Journal of Educational Research

Official Organ of the National Association of Directors of Educational Research

B. R. BUCKINGHAM, *Editor*

Associate Editors

E. J. Ashbaugh — Leonard P. Ayres — W. W. Charters — S. A. Courtis — Walter S. Monroe — George D. Strayer — Lewis M. Terman

VOL. II SEPTEMBER, 1920 No. 2

CONTENTS

A Testing Program for Elementary Schools, *B. R. Buckingham and Walter S. Monroe* 521

The Use of Psychological Tests in the Vocational Guidance of High-School Pupils, *W. M. Proctor* 533

Standardizing Procedure in a Combination System of Supervised Study, Varying Scope of Work, and Weighted Credit, *W. H. Hughes* 547

Measurements in Fairy Tale, *Cyrus D. Mead* 557

Provisions for Individual Differences in Teaching Reading, *W. W. Theisen* 560

Scale of Attainment No. I—An Examination of Achievement in the Second Grade, *Luella W. Pressey* 572

Editorials 582

Reviews and Abstracts 587

News Items and Communications 592

National Association of Directors of Educational Research 598

Entered as second-class matter January 13, 1920, at the Post Office at Menasha, Wisconsin, under the act of March 3, 1879.

Journal of Educational Research

Copyright 1920, by
Public School Publishing Co.

Published for the Bureau of Educational Research, University of Illinois

Editor-in-Chief: **B. R. Buckingham**
Director of Bureau of Educational Research, University of Illinois

Associate Editors

E. J. ASHBAUGH, *Ex-Officio, Secretary of Research Association, University of Iowa, Iowa City*
LEONARD P. AYRES,
Ed'l Sec'y, Russell Sage Foundation, N. Y.
W. W. CHARTERS,
Carnegie Inst. Tech., Pittsburgh
S. A. COURTIS,
City Schools, Detroit, Mich.
WALTER S. MONROE,
University of Illinois, Urbana, Ill.
GEORGE D. STRAYER,
Columbia University, New York City
LEWIS M. TERMAN,
Leland Stanford Junior University, Stanford University, Calif.

Contributing Editors

VIRGIL E. DICKSON,
Director of Research, Public Schools Oakland, California.
CHARLES FORDYCE,
Dean Teachers' College, and Director of Bureau of Educational Measurements, University of Nebraska.
WILLIAM S. GRAY,
Dean College of Education, University of Chicago.
M. E. HAGGERTY,
Director Bureau of Co-operative Research, University of Minnesota.
V. A. C. HENMON,
Director School of Education, University of Wisconsin.
ARTHUR W. KALLOM,
Assistant Director Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.
F. J. KELLY,
Dean, School of Education and Director Bureau of School Service, University of Kansas.
ROBT. H. LANE,
Assistant Superintendent of Schools, Los Angeles, Calif.
WILLIAM A. McCALL,
Assistant Professor of Education, Teachers' College, Columbia University.
GEORGE W. MELCHER,
Director Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri.
EUGENE A. NIFENECKER,
Assistant Director Bureau of Reference, Research, and Statistics, Public Schools, New York City.
JOSEPH P. O'HERN,
Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York.
LEON O. SMITH,
Assistant Superintendent (in charge of Educational Research) Public Schools, Omaha, Nebraska.
CLARENCE W. SUTTON,
Director Division of Reference and Research, Public Schools, Cleveland, Ohio.
W. W. THEISEN,
Supervisor of Educational Measurements, Public Schools, Cleveland, Ohio.
CLIFFORD WOODY,
Assistant Professor of Education, University of Washington.

THE JOURNAL OF EDUCATIONAL RESEARCH is published at 450 Ahnaip St., Menasha, Wis. monthly except July and August by Geo. Banta, Official Printer to the Public School Publishing Co. The subscription price is $4.00; $4.30 in Canada or Mexico; and $4.50 in other foreign countries.

The editorial office is at Urbana, Illinois. Address all communications on educational matters, articles for the magazine, etc. to B. R. Buckingham, University of Illinois, Urbana, Illinois.

The JOURNAL OF EDUCATIONAL RESEARCH is both needed by and needs the subscriptions of many more teachers and school officers.

The title page and index for Volume I will be furnished with the September, 1920, number.

The Public School Publishing Company conducts the subscription work. Send all orders to them at Bloomington, Illinois.

Articles to Appear Soon

Selection of Silent Reading Textbooks by *Ernest Horn.*

Value of a Handwriting Scale to an Untrained Teacher by *R. L. Morton.*

Reliability of the Binet Scale and Pedagogical Scales by *Arthur S. Otis* and *Herbert E. Knollin.*

Reliability of Test Scores by *Truman L. Kelley.*

Journal of Educational Research

Editorial Office, Urbana, Illinois — Editor-in-chief B. R. Buckingham, University of Illinois

VOLUME II — SEPTEMBER, 1920 — NUMBER 2

A TESTING PROGRAM FOR ELEMENTARY SCHOOLS

B. R. BUCKINGHAM AND WALTER S. MONROE
University of Illinois

Most of us are agreed that tests, in order to produce the best results, must be given by teachers. Only in a few school systems can they be administered either by the superintendent in person or by some one acting for him. Simple tests, therefore, will be in greatest demand—tests which are easily administered, objectively rated, and quickly interpreted. Monroe's Standardized Silent Reading Test and Courtis' Standard Research Tests in Arithmetic have been widely used largely because they possess these characteristics of simplicity. The test-maker's art is most clearly revealed when he is able to devise tests which are not only simple but valid as measures of significant school abilities.

The superintendent, therefore, who contemplates initiating a testing program, or even the continuance of one already begun but not carried forward to a point where teachers have become relatively expert, will do well not to be too ambitious in the selection of the number and character of the measuring instruments which he will use. In a certain sense he will have to "sell" the idea to his teachers. He cannot do this if he imposes upon them a confusing and difficult task. It is better for him to select a few tests, each concerned with important activities, than to try to cover a wide range of school subjects. One or two tests carefully administered and competently followed up will satisfy and convince; many tests may do no more than confuse and antagonize.

This would seem to be particularly true in the case of rural schools. When, therefore, the Illinois Association of County Superintendents appointed a "Committee on Standard Tests" and when this committee, pursuant to instructions, met with us to plan a rural-school testing program, we were at once convinced

that such a testing program would have to be simple. Even if general considerations had not dictated such a policy, the relative inexperience and lack of training of rural-school teachers in administering tests would have convinced us that we should have to be content with small beginnings.

At the same time we recognized that the opportunity was unusual. The testing of rural schools on any considerable scale has not been attempted. Their isolation prevents, except at prohibitive cost, the employment of special examiners. Quite generally the teachers have no experience in giving tests; and the long-range supervision of the county superintendent necessarily precludes the careful oversight which alone can surmount these obstacles. More favorable conditions have for the most part confined the test movement to city systems. Yet it is unquestionably true not only that the rural schools ought to be tested, but that they are in a position to profit by the testing to a peculiar degree. The mere fact that the grades are all under one teacher facilitates any regrouping of children which the results of testing may suggest. Moreover, these grades are, or may be, loosely organized. A child may easily—far more easily than in a graded school—be instructed with one grade in one subject and with another grade in another subject. For these reasons and for many others we were eager to meet the county superintendents in their proposal to test the rural schools of the state.

When the committee met with us, we proposed first, that not more than two educational tests should be used; second, that these tests should be in the most important subjects—i.e., reading and arithmetic; third, that if it were in any way possible, the intelligence of the children should be measured; and fourth, that the pupils' scores in the educational tests on the one hand and in the intelligence test on the other should be definitely related to each other. To these suggestions the committee readily agreed.

But what reading test, what arithmetic, and, above all, what intelligence test should we use? As to the reading test there was not much chance for argument. If we were going to have something simple, we should clearly want to use Monroe's Standardized Silent Reading Test or some test of that character. In arithmetic we first considered Courtis' Standard Research Test, Series B; but we afterwards rejected it in favor of two series of

tests, one for grades III, IV, and V, and the other for grades VI, VII, and VIII. This action was taken because of the varied types of examples which such a plan would provide. The selection of an intelligence test constituted the chief problem. Although we now believe that our rural-school program of testing has a wide and general usefulness, it is well to point out that in our relations with the county superintendents our attention was narrowed to the particular project in hand. Since, therefore, the money available for testing in each county was limited, we were obliged to select an inexpensive intelligence test or face the fact that very few superintendents would be able to undertake our program. But no cheap intelligence test seemed to be good enough for our purpose. In fact, we were confronted by the condition—common enough in these days—that the thing which was good enough to be desirable was also too expensive to be had.

The only way out of the difficulty seemed to be the devising of a new intelligence test. Of course, the expense of such a procedure, when combined with that of printing and distributing, might well be greater than the expense of purchasing tests already on the market, but the cost of devising the intelligence test would be a legitimate part of our cooperation with the superintendents. Our final proposal was, therefore, that we would make up a battery of tests in reading, arithmetic, and intelligence, and that we would print all the material in one booklet. The advantages of bringing the material together in this way hardly need be pointed out. By this means a pupil's record on all the tests would be kept together, and the calculation of certain derived measures (to be explained later) would be facilitated.

This battery of tests we agreed to devise, standardize, print, and distribute in quantities sufficient for the use of the rural schools of the state, and to have it ready for them when they opened in September. Moreover, we agreed to prepare a handbook of instructions to teachers, and to have it printed in time for use at a county superintendents' convention to be held at the University of Illinois on the twenty-ninth and thirtieth of July.

This entire project was rendered especially difficult because the school year was already far advanced. The Committee on Standard Tests first met with us on April 2, 1920. The decision to devise and print a testing instrument for use in the rural schools

was not reached until about ten days later; and the time required for printing the tests and for delivering them to superintendents—delayed as delivery was by railroad strikes—carried us well into May. Even with plenty of time at our disposal, the making of a new intelligence test would not have been a task lightly to be undertaken. Such work involves careful investigation of existing material, the formulation of preliminary editions, a try-out on thousands of children, and endless statistical work, typing, mimeographing, and printing. Nevertheless, such a course of action seemed necessary in order to save the situation.

By making a special effort, and with the effective cooperation of a number of Illinois city superin endents, we were able to carry out our plans. Of the intelligence test we used two preliminary editions and of the reading test one. The arithmetic test is a composite of material of known difficulty mostly from Monroe's Diagnostic Test. The final edition of the test booklet is now coming from the press. It consists of 200,000 copies and over 100,000 of them have been spoken for. The teacher's handbook is likewise ready.

In the compilation of the "Illinois Examination," as we have called this group of tests, we do not claim to have displayed any particular originality. Certain materials both for intelligence and educational testing are to be had for the asking, and such material we have used. In this respect we resemble Homer as represented by Kipling in his introduction to the "Barrack-room Ballads":

> When 'Omer smote 'is bloomin' lyre,
> He'd 'eard men sing by land an' sea;
> An' what he thought 'e might require,
> 'E went an' took—the same as me!

Even if time had permitted us to originate all the material of which the tests are composed, it is doubtful whether our time would thus have been most serviceably employed. At any rate, we have preferred to be useful first and original afterwards.[1]

The "Examination" appears in two parts, and is accompanied by record sheets and the handbook of instructions to which we

[1] Dr. Sidney L. Pressey kindly permitted us to use his tests of Verbal Ingenuity and Arithmetical Ingenuity. Professor E. H. Cameron contributed the Substitution Test.

have referred. Part I is for grades III, IV, and V. It comprises the following material:

1. *The Intelligence Test*, consisting of sub-tests in analogies, arithmetical problems, sentence vocabulary, substitution, verbal ingenuity, arithmetical ingenuity, and synonym-antonym.
2. *A Revision of Monroe's Standardized Silent Reading Test for Grades* III, IV, *and* V.
3. *A Test of Ability in the Operations of Arithmetic*, consisting of eight sub-tests—four on the combinations, and four involving simple examples in each operation.

Part II is for grades VI, VII, and VIII. The intelligence test is the same as in Part I. The reading test is Dr. Monroe's revision of his Standardized Silent Reading Test for grades VI, VII, and VIII. Although this material differs from the reading material for Part I, the instructions are in every way the same. The arithmetic test consists of seven sub-tests as follows: column addition, long multiplication, long division, subtraction, addition and subtraction of fractions, multiplication and division of fractions, division of decimals.

The silent reading test yields measures of both rate and comprehension. For the operations of arithmetic, it was considered desirable, however, to have a single score which would serve as an index of the pupil's achievements. This was secured merely by taking as the pupil's score on each test the number of examples right. These scores were then added after certain adjustments had been made to equalize the weight given to the several tests. The sum thus obtained was the pupil's point score on the operations of arithmetic.

The actual working time of pupils is 27 minutes for Part I and 37½ minutes for Part II. The total time required for giving either of the two parts is about the same—namely, one hour. It is recommended that Part I be given in two separate periods, but that these periods be included in the same half-day session. This is clearly advisable because all the pupils who are present at one test period will then be present at the other.

It is clear that in a rural school (or in a graded school in which children of the fifth and sixth grades meet in the same room) both parts of the examination will have to be administered by the same teacher. This, however, does not mean that two hours of the teacher's time need be devoted to the task. The same preliminary

statement to children and the same intelligence test are used for all grades. Moreover, although the material in the reading tests differs for the two groups of grades, the instructions are identical, and all children may, therefore, be tested simultaneously. Accordingly, the teacher who is giving both parts of the examination may administer them together so far as the intelligence and reading tests are concerned.

The union of intelligence and subject-matter testing has obvious advantages; and these advantages will naturally be more surely obtained if, as in this instance, the intelligence and subject-matter tests are mechanically combined—i.e., form parts of the same testing instrument. But even then the advantages cannot be fully realized unless some easy way is found of combining the scores on the several tests. If the Illinois Examination presented no characteristics other than those we have mentioned above, we should hardly be justified in devoting this article to it. It is, however, in the treatment of scores that the examination is truly distinctive.

The sum of the scores obtained by a pupil on the sub-tests of the intelligence scale yields his "point score" for that scale. On the basis of our returns, we have ascertained for each half-year of age the median point score on the intelligence test. The point score of a pupil may then be interpreted with reference to these medians. For example, it is found that 97 is the median point score of children 14 years of age; accordingly, a child whose score is 97 indicates by his performance a mentality which is typical of that age. Fourteen may then be said to be his "mental age."

This conception of mental age is valuable—much more so than that of chronological age. It indicates the degree of mental development which the child has attained at the time the test is given. Since school work addresses itself primarily to mentality, it is, so to speak, the efficient age with reference to school work. Indeed, it may without much inaccuracy be defined as the present ability to learn. It is, therefore, far more important as a basis of action than the number of years a child has lived.

In childhood, however, mental age like chronological age is constantly changing; but unlike chronological age, its change is not registered merely by the flight of time. For a basis of action which contemplates a continuing educational program, we need

some determination (derived if possible from the point score) which shall be relatively unchanging over a period of years. It has been found that one form of relationship between a child's mental age and his chronological age fulfills this condition. This relationship is expressed by dividing the mental age by the chronological age, or stated another way, by finding the percent that the mental age is of the chronological age. The result is called the "intelligence quotient" (I. Q.). Thus, if a child's mental age as determined from his point score is 9 years and his chronological age is 12 years, his I. Q. is 75.[2] This expresses the fact that during the 12 years that this child has lived his annual mental development has been but 75 percent of normal. In other words, each elapsed year of his life he has developed mentally but three-quarters of a year. On the other hand, if this child of 12 had a mental age of 15, his I. Q. would be 125. This would indicate an annual mental development amounting to one and one-fourth years. An intelligence quotient of 100 arises from the fact that mental and chronological ages are the same. It shows that mental development is normal. Moreover, since the I. Q. is relatively constant, these percents of development will continue to manifest themselves. This fact affords a basis of prophecy. We may be sure, for example, that certain children cannot advance beyond the sixth or seventh grade; and they should, therefore, have a modified course of study with reference to that fact. On the other hand, we may be reasonably confident that certain children can, if we let them, finish the elementary school while they are still young. It is our duty to plan for them a richer curriculum and an accelerated progress. If their intelligence quotients indicate that they are advancing mentally at a more than normal rate, then they should be advanced in school at a corresponding rate.

In order to make it unnecessary for teachers, when using the Illinois Examination, to compute mental ages and I. Q.'s, we have provided as part of the accessory material a table of which Table I is a section. Its use may be illustrated by supposing that an eleven-year-old child achieves a point score of 57. Note that the entry "57" in the first column is written opposite "10–0" in the second column. This indicates that the pupil in question has a mental age of 10 years. Since the child is 11 years old, note also the column under "Chronological Age" which is headed "11–0."

[2] Conventionally, the quotient is multiplied by 100 to get rid of the decimal point.

By tracing this column to a position opposite the point score of "57," the entry "91" is found. This is the I. Q. of the child in question.

TABLE I. INTELLIGENCE QUOTIENTS

Point Score	Mental Age	Chronological Age				
		10-0	10-6	11-0	11-6	12-0
82	12-6	124	119	113	109	104
77	12-0	120	114	109	104	100
72	11-6	115	109	104	100	96
67	11-0	110	105	100	96	92
62	10-6	105	100	96	91	88
57	10-0	100	95	91	87	84
52	9-6	95	91	87	83	80
47	9-0	90	86	82	79	76

The most distinctive characteristic of educational tests has been the fact that they have been standardized. It has not, however, been generally recognized that our present standards apply only to the median or average scores of a group (usually a grade) and not to the scores of individual pupils. It is true that individual scores may be compared with our present grade standards, but only very limited conclusions may be drawn. For example, the fifth-grade standard for Monroe's Standardized Silent Reading Test for comprehension is 21. The scores of a typical fifth-grade class will range from about 10 to 30. If the median score is 21, we say that the *class as a whole* is up to standard although half of the pupils are above the grade standard and the other half are below it. We do not know whether any particular pupil has done as well as he should or not. A pupil may fail to achieve as high a comprehension score as his intelligence (mental age) indicates he should, and at the same time he may have a score above the grade standard. In other words, a bright pupil may be "above standard for his grade" but below *his own* standard. The opposite also is true. A pupil who has a score "below standard for his grade" may have made a very high score in comparison with his intelligence.

The need for individual standards to use in interpreting the scores of individual pupils is particularly urgent when educational tests are used in rural schools because the number of pupils in any one grade is generally so small that it cannot be representative. Not infrequently there are only two or three pupils. In such cases the median score can have only a very limited signif-

icance. This makes the interpretation of the scores of individual pupils imperative and for doing this individual standards are necessary.

A standard score is the score which a pupil should make on a given test. In the case of silent reading and the operations of arithmetic, two obvious factors, upon which the score which a pupil should make depends, are his intelligence or capacity to learn and the amount of his training or roughly the school grade he has attained. The data obtained by giving both the intelligence test and Monroe's Standardized Silent Reading Test to several thousand children indicated that the amount of instruction which the pupil had received contributed very little to the score which he made on the reading test. It was found that the average silent reading score is approximately the same for pupils of a given mental age whether they are in the third grade or the eighth. For this reason it was deemed appropriate to standardize the educational tests of the Illinois Examination on the basis of mental age. Accordingly, the scores of all pupils of each mental age were assembled and the median taken as standard.[3] In this way, the standard for each half-year of mental age was obtained.

A pupil's point score may now be compared with *his own* standard, i.e., the standard for his mental age. However, it will be more convenient, and will also make possible some other things, if the mental age standards are used to translate a pupil's point score into an *achievement age score* (A. A.). This may be done in the same way as point scores on the intelligence scale were translated into mental ages. When a pupil's score on an educational test (silent reading or the operations of arithmetic) is expressed as an achievement age score, his mental age becomes his standard. Since both his score (A. A.) and his standard (M. A.) are expressed in terms of a common unit, they may be compared by dividing in the same way that his mental age was divided by his chronological age to obtain his intelligence quotient (I. Q.). The quotient of A. A. divided by M. A. we have called *achievement quotient* (A. Q.).

Table II illustrates, for a brief range of point scores and mental ages, the device which we have provided for securing these derived measures. The first line of Table II may be read as follows:

[3] In doing this it was recognized that the educational tests for grades VI, VII, and VIII were more difficult than the ones for the lower grades, and corrections were made in order to adjust the scores of the more difficult tests to the scale of the easier ones.

A pupil whose rate score in reading is 155, whose comprehension score is 12, and whose arithmetic score is 63 shows on all three of these counts an achievement age of 12 years and 6 months. His achievement quotient for each of the three scores is 124 if his mental age is 10 years, 119 if his mental age is 10 years and 6 months, 113 if his mental age is 11 years, *et cetera*.

TABLE II. ACHIEVEMENT QUOTIENTS

Point Scores			Achievement Age	Mental Age				
Rate	Comprehension	Arithmetic		10-0	10-6	11-0	11-6	12-0
155	12	63	12-6	124	119	113	109	104
150		60	12-0	120	114	109	104	100
145	11	57	11-6	115	109	104	100	96
140		53	11-0	110	105	100	96	92
135	10	49	10-6	105	100	96	91	88
130		44	10-0	100	95	91	87	84
125	9	40	9-6	95	91	87	83	80
120	8	35	9-0	90	86	82	79	76
114		30	8-6	85	81	78	74	71
107	7	25	8-0	80	77	73	70	67

An achievement quotient of 100 means that the pupil has achieved exactly as well as the average of the pupils of his mental age. If his achievement quotient is 130, it means that he has achieved 30 percent more than the average of the pupils of his mental age; on the other hand, if his achievement quotient is 75, we have evidence that he has achieved only 75 percent as much as the average of pupils of his mental age.

The value of individual standards and the achievement quotient is illustrated in Figure 1. This figure shows the scores for comprehension and the achievement quotients of pupils in a fifth-grade class. The former are plotted along the horizontal axis and the latter along the vertical axis. Each pupil is represented by a dot. Distances from the two axes show the two measures of the pupil's achievement. The grade standard in terms of a point score is indicated by the arrow.

The lowest score in the class is four. If we had only the grade standard, all we could say about this pupil would be that he is below standard and at the foot of his class. His A. Q. shows that in comparison with his own standard he has achieved more than is usually achieved by a pupil of his mentality. In fact

when his mental age is considered, he is one of the "good" pupils in his class.

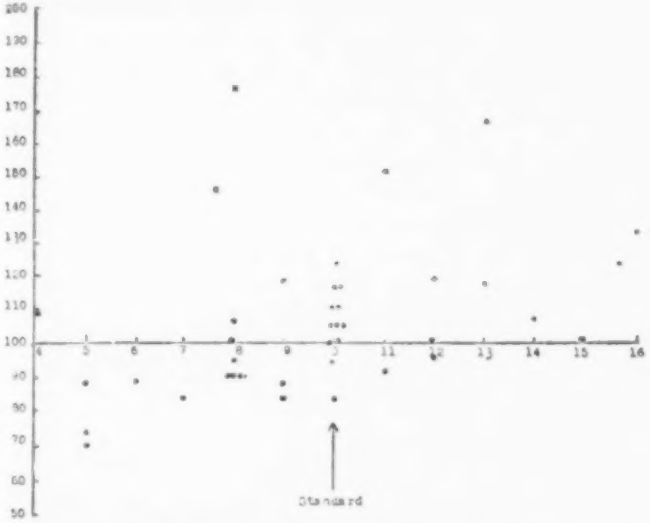


FIGURE 1. SILENT READING COMPREHENSION. RELATION BETWEEN POINT SCORES AND ACHIEVEMENT QUOTIENTS. GRADE V

An added advantage of this plan is that when we have transmuted point scores in subject-matter tests into achievement ages, we have reduced them all to the same units in the sense that each successive year corresponds to an increment of ability gained by typical children in equal lengths of time. Since the achievement ages are expressed in the same units, they may be combined as point scores cannot be combined. For example, we have no obvious way of expressing the total achievement of a child who scores 155 in rate, 11 in comprehension, and 49 in arithmetic. By reference to Table II, however, we observe that these point scores indicate respectively achievement ages of $12\frac{1}{2}$, $11\frac{1}{2}$, and $10\frac{1}{2}$ years. On the assumption that the abilities to which the point scores refer are all equally important, we may obtain the simple mean of these ages, and we may thus express the composite achievement age as $11\frac{1}{2}$ years.

Clearly, the relation of achievement age to mental age is important. If a pupil's achievement age is more than his mental age, it means that some factors other than intelligence are favorable.

The school is doubtless doing its work well. Of course, there are innate characteristics other than intelligence which may be operative. The most important of these is effort; but since any effort greater than usual is not likely to be elicited except through the superior work of the school, it is probably not inappropriate for the school to take to itself credit for the showing.

On the other hand, if achievement age is less than mental age, the same course of reasoning leads us to suspect that the school is at fault. Of course, this may not be true. For example, irregular attendance for reasons over which the school has no control might produce this result. At any rate, when achievement age is less than mental age—i.e., when the achievement quotient is less than 100—an investigation should be made.

Even if the combination of an intelligence test with subject-matter tests were not necessary to secure individual standards, we believe that any competent testing program should involve both kinds of tests. When the latter are given alone and results are compared with standards, great injustice may be done to the school and to the teacher. If the results are below standard, the teacher and the school are likely to be criticized. This may be wholly unjust. The general level of mentality among the children may be below normal by an even greater amount than the scores in subject-matter tests are below standard. On the other hand, scores above standard in educational tests may not indicate that the school is doing good work because intelligence may be above standard to an even greater degree.

Again, the use of intelligence tests without subject-matter tests is one-sided. The real truth about a pupil or a class is made evident when to a knowledge of mentality is added a knowledge of achievement. If mental age may be taken as representing potential ability, achievement age may be taken as representing *effective* ability. The one shows a child's chance to learn, the other what he *has* learned; the one measures his ability as a possession, the other his ability as put to use. Since achievement age indicates ability as it functions, it affords an important supplement to, and commentary upon, mental age as a basis of action—whether that action involves the reclassification of pupils or some other modification of school procedure. Moreover, it affords, when related to mental age, a stimulus to effort or a basis of satisfaction according as it is less or more than the mental age.

THE USE OF PSYCHOLOGICAL TESTS IN THE VOCATIONAL GUIDANCE OF HIGH-SCHOOL PUPILS

W. M. PROCTOR
Leland Stanford Junior University

The use of psychological tests as a basis for the prediction of probable school success,[1] and in the educational guidance[2] of high-school pupils has been discussed by the writer in previous papers in this Journal. The use of such tests in the vocational guidance of high-school pupils will be discussed in this paper.

THE EMPLOYMENT MANAGER AND THE VOCATIONAL COUNSELOR

The application of psychological tests to the selection of men for particular jobs is well along in the experimental stage.[3] The employment manager has demonstrated the value of such tests when used in the selection of employees. If the tests are carefully devised to discover whether the applicants have certain essential mental qualifications for the kind of work they will be called upon to perform, and if final selection is made from the few who make high scores in the tests, the results are likely to be satisfactory to the employer. When in addition to specific mental tests, appropriate trade or performance tests are given, the employment manager or personnel expert can select from the highest 25 percent of applicants with reasonable assurance that they will make good at the tasks assigned to them. These tests tend effectively to sift out the undesirables, to reduce the amount of labor turnover, and to secure for the employer a higher grade of employees.

The person who selects men for a particular occupation needs only to know the qualifications for success in that occupation. He can plan his tests with a view to eliminating all those who do not measure up to the established standard. If only five out of one hundred applicants are selected and they all prove to be

[1] Proctor, W. M. "Psychological tests as a means of measuring the probable school success of high-school pupils," *Journal of Educational Research*, 1:258–70, April, 1920.

[2] Proctor, W. M. "The use of psychological tests in the educational guidance of high-school pupils," *Journal of Educational Research*, 1:369–81, May, 1920.

[3] Link, H. C. *Employment psychology*. New York: Macmillan Co., 1919.

adapted to their work the tests by which they were selected are counted as satisfactory. The ninety-five rejected applicants do not concern the employment manager.

The vocational counselor, however, has to think of the ninety-five as well as the five. His field is a broad one. He is called upon to advise people possessing every variety of fitness to enter every possible kind of occupation. It would be manifestly impossible for the vocational counselor to give adequate trade or psychological tests corresponding to the infinite variety of occupations open to American youths.

It is true that he must avail himself of every possible scientific aid in arriving at his conclusions. His preparation will necessarily include a wide knowledge of occupations, and special training in the discovery of occupational aptitudes. But he should never persuade himself or lead others to believe that he is able to chart unerringly their abilities and give them an absolute vocational classification. The successful counselor will understand at the outset that he is a guide and not a dictator, and that he is dealing with probabilities and not with certainties.[4]

Occupational Levels of Intelligence

Among the factors which the counselor must take into account in estimating the probability of a person's success in an occupation, the intelligence level is one of the most important. If we accept Stern's definition of intelligence as a working basis for the discussion of the subject, there can be no question of the vital connection between intelligence and vocational success. Stern says:[5] "Intelligence is a general capacity of an individual consciously to adjust his thinking to new requirements: it is general mental adaptability to new problems and conditions of life." Given a high degree "of mental adaptability to new problems and conditions of life" coupled with energy, persistence, and reliability, it would be reasonable to assume that one's range of possible vocational success would be wide. Conversely given a low level of intelligence, even though the other qualities mentioned are present, one's range of possible vocational success would be greatly restricted.

[4] Kitson, H. D., "Vocational guidance and the theory of probability," *School Review* 28:143-50, February, 1920.

[5] Stern, William. *The psychological methods of testing intelligence.* Baltimore, Warwick & York, 1914, p. 3.

We have already gone far enough in the development of intelligence tests to justify the statement that individual levels of intelligence can be discovered with approximate accuracy. The next step, so far as vocational guidance is concerned, is to discover whether or not there are occupational intelligence levels, i.e. levels of intelligence more or less characteristic of the workers in a given occupation. That there are discoverable differences in the intelligence levels of workers in the various occupations is suggested by the findings of the Division of Psychology, Sanitary Corps, United States Army.[6]

Table I is derived from the chart found on page 23 of the pamphlet, *Army Mental Tests*. The median ratings of this chart are changed from the letters "A," "B," etc. to the raw scores on the Alpha Army Scale (which range from 0 to 212). For each of the 43 occupations selected from the 72 shown on the chart, the median score is given, as well as the range of the middle 50 percent of the scores. The top line of the table would then read: laborers, median score, 35; range of scores made by the middle 50 percent, 21–63. It should then be understood that 25 percent of the laborers scored less than 21 points, while the top 25 percent scored over 63 points out of a possible 212 points. The chart was made up from the returns of approximately 36,500 men, and the data were taken from the soldiers' qualification cards.

Figure 1 illustrates graphically the spread of the middle 50 percent of the scores on Army Alpha by occupational groups. There is considerable overlapping. The unskilled, semi-skilled, and skilled labor groups differ but little as to the beginning of the middle 50 percent of scores (21, 23, and 26 respectively). But there is a distinct difference in the upper limits, which are 63, 70, and 95 respectively. The beginning of the middle 50 percent of the business and clerical group is nearly as high as the upper limit for the unskilled labor group, and the beginning of the middle 50 percent of the professional group is higher than the upper limit of the skilled labor group.

If the scores of the entire number of men examined by the Division of Psychology, Sanitary Corps, of the United States

[6] *Army mental tests methods, typical results, and practical applications.* Washington: Government Printing Office, November, 1918.

TABLE I. OCCUPATIONAL INTELLIGENCE LEVELS, BASED ON ARMY PSYCHOLOGICAL TESTS OF 36,500 MEN. ALPHA SCALE

Occupations	Median Score	Range of Middle 50 Percent
LABORERS (UNSKILLED)	35	21 to 63
SEMI-SKILLED LABOR		
Cobblers	39	23 to 67
Teamsters	41	23 to 68
Farm workers	42	24 to 70
Barbers	43	22 to 70
Horse-shoers	44	25 to 70
SKILLED LABOR		
R. R. shop-mechanics	45	26 to 83
Bricklayers	48	23 to 81
Cooks	49	28 to 79
Bakers	53	35 to 83
Painters	53	31 to 79
Blacksmiths	54	29 to 83
Bridge-carpenters	55	27 to 84
General carpenters	57	33 to 85
Butchers	58	33 to 85
Locomotive enginemen	59	33 to 82
Machinists	61	33 to 86
R. R. conductors	62	40 to 84
Plumbers	62	38 to 87
Tool-makers	63	41 to 88
Auto-repairmen	63	41 to 89
Chauffeurs	63	38 to 90
Tool-room-experts	64	43 to 88
Policemen-detectives	64	44 to 89
Auto-assemblers	65	44 to 97
Ship-carpenters	66	49 to 95
BUSINESS AND CLERICAL		
Telephone operators	70	58 to 99
Concrete const'n foremen	75	48 to 116
Photographer	77	52 to 104
General electrician	82	58 to 110
Telegraphers	84	59 to 107
R. R. clerks	92	66 to 116
General clerks	96	74 to 123
Mechanical engineers	98	63 to 133
Bookkeepers	99	78 to 126
Dental officers	106	84 to 130
Mechanical draughtsmen	112	79 to 134
Stenographers	115	93 to 142
Accountants	117	101 to 145
PROFESSIONAL		
Civil engineers	125	98 to 147
Medical officers	130	101 to 165
Army chaplains	150	109 to 173
Engineer officers	157	134 to 184

Approximate occupational intelligence levels:

1. Unskilled labor	Median	35,	Middle	50%	21	to	63
2. Semi-skilled	"	42,	"	"	23	to	70
3. Skilled labor	"	61,	"	"	26	to	95
4. Business and clerical	"	96,	"	"	58	to	145
5. Professional	"	140,	"	"	98	to	184

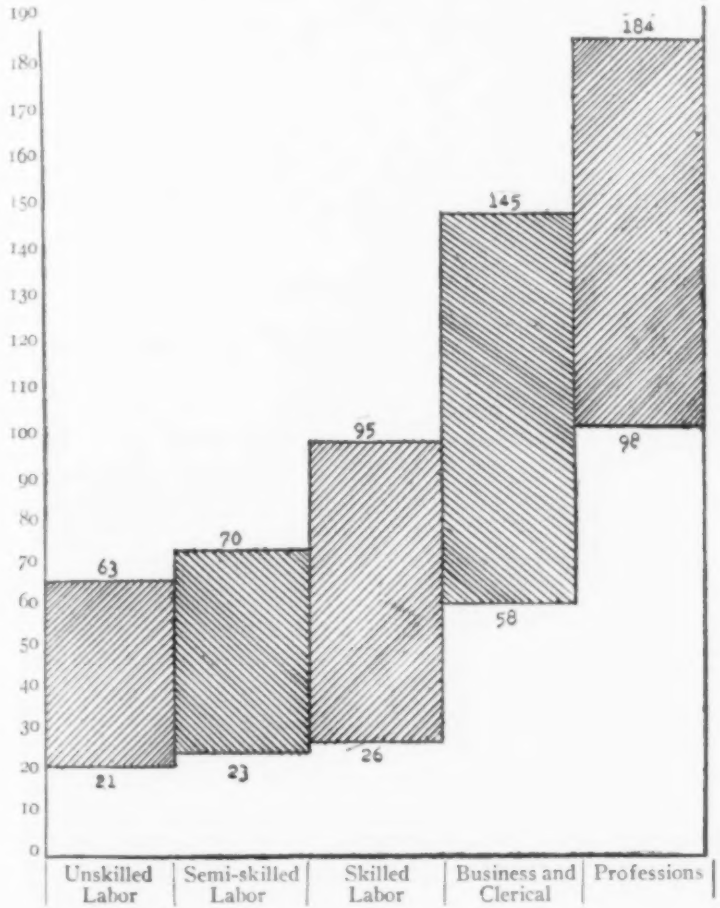


FIGURE 1. SHOWING SPREAD OF MIDDLE 50 PERCENT OF CASES, TABLE I, BY OCCUPATIONS

Army, could be grouped by occupations the final results of compilation would probably show some deviations from the medians and middle 50 percents given in Table I. But the general tendencies therein indicated would no doubt be confirmed. There would be found more or less clearly defined levels of intelligence in the various occupations, corresponding roughly to the amount of intelligence necessary to succeed in them. There would be much overlapping and within each occupation a wide range of intelligence would be found. But in the occupational groups above unskilled labor one would expect to find critical scores[7] or points below which occupational success could not be expected.[8] In the professional group for example, one would expect to find the greatest number of occupational failures among the lowest 25 percent, i.e. among those who made scores ranging from 0 to 98.

Application to Guidance of High-School Pupils

How a knowledge of individual and occupational intelligence levels may be utilized in the vocational guidance of high-school pupils can be illustrated by the data presented in Table II. In this table 930 pupils in eight high schools are distributed according to vocational ambition and scores made on Army Scale Alpha and Examinations *a* and *b*. The different occupational choices have been divided into five groups: agriculture, mechanical and industrial, business and clerical, unclassified,[9] and professional. The professional group covers 50.5 percent of the choices, 470 cases; unclassified, 11.2 percent, 104 cases; business and clerical, 29.7 percent, 276 cases; mechanical and industrial, 3.6 percent, 34 cases; and agricultural, 5.0 percent, 46 cases.

The need for vocational guidance of high-school pupils is brought out very clearly by the way in which the choices are concentrated in the professional, semi-professional (i.e. unclassified), and business and clerical divisions of Table II. These include 31 different occupations, popularly known as "white collar jobs," and comprise 91.2 percent of the 930 choices. Agri-

[7] Thurstone, L. L. "Mental tests for college entrance," *Journal of Educational Psychology*, 10:129–41, March, 1919.

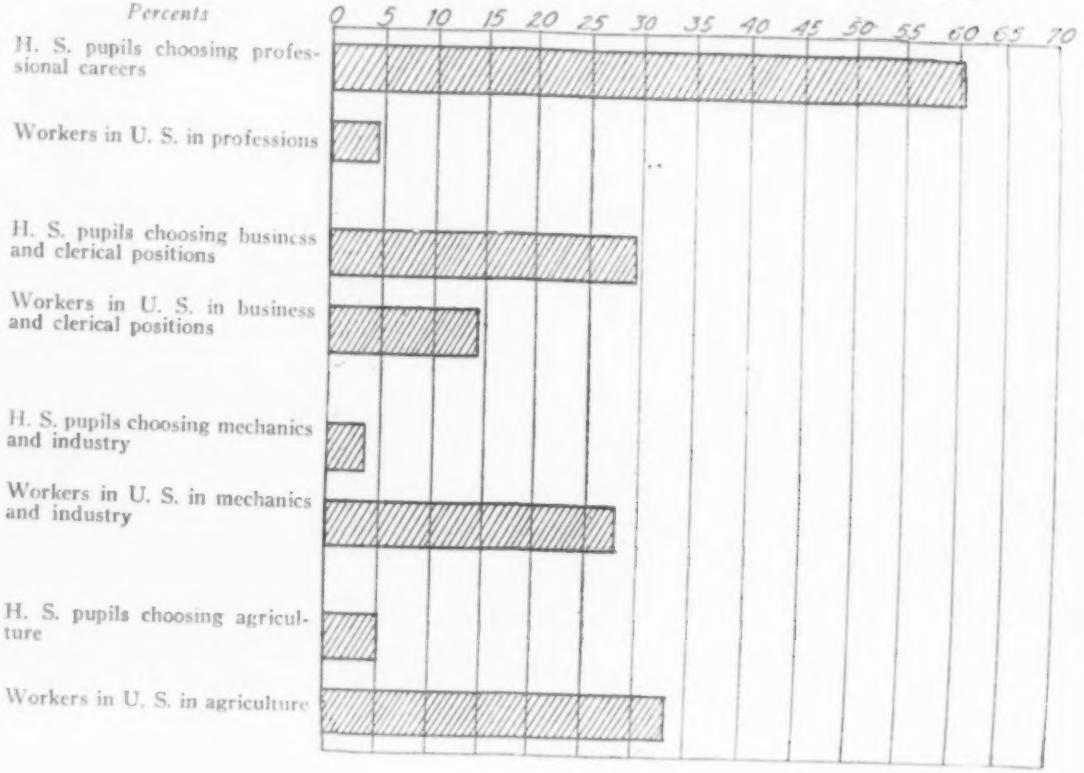
[8] Cowdery, K. M. "A statistical study of intelligence as a factor in vocational success," *Journal of Delinquency*, 4:227, November, 1919.

[9] Most of the occupations belonging to this group are called "professional" by the United States Census Bureau.

TABLE II. VOCATIONAL AMBITIONS OF 930 HIGH-SCHOOL PUPILS DISTRIBUTED ACCORDING TO INTELLIGENCE RATINGS ON ARMY SCALE ALPHA AND GROUP EXAMINATIONS *a* AND *b*

Letter rating	C—	C		C+		B		A	Total	
Alpha	25–44	45–59	60–74	75–89	90–104	105–119	120–134	135–212		
a and *b*	25–49	50–67	68–84	85–99	100–119	120–137	138–154	155–237		
Choice of Occupation										
AGRICULTURE										
Agriculture		2	1	5	4	7	10	12	41	
Floriculture					1				1	
Forestry			1			1		2	4	46
MECHANICAL AND INDUSTRIAL										
Auto-mechanic					4				8	
Dressmaker						2			2	
Draughtsman			2	5		3	8	3	21	
Electrician						1			1	
Contractor					1				1	
Milliner					1				1	
Printer						1	1		2	
Aviator					1		1		2	34
BUSINESS AND CLERICAL										
Banker				1				1	2	
Bookkeeper		3	5	10	22	9	7	6	62	
Business manager		1	2	6	5	3	1	5	23	
Civil Service							1		1	
Clerk (sales)			1	1		1			3	
Manufacturer					1	1			2	
Real estate				1		1		1	3	
Salesman (Tr)					2	1	1		4	
Stenographer	2	2	11	24	41	38	33	25	176	276
UNCLASSIFIED										
Actress				2	1	1			4	
Army and Navy					3	3	2	6	14	
Artistic career			2	2	10	5	6	3	28	
Dancer							1		1	
Dog Fancier					1				1	
Home-maker			1		1	1	1	1	5	
Nurse			4	7	10	11	8	6	46	
Sea-captain					1				1	
Social Service						1		1	2	
Writer							1	1	2	104
PROFESSIONAL										
Architect						3			3	
Bacteriologist							1	1	2	
Chemist				1		3	3	7	14	
Dentist		1		2	3	1	2	2	11	
Druggist						2		3	5	
Engineering			2	12	20	30	32	44	140	
Journalism					2	3		5	10	
Law					2	2	9	7	20	
Librarian						1	2	2	5	
Ministry							1		1	
Physician				1	4	11	10	10	36	
Teaching			8	23	44	68	47	33	223	470
Totals	2	9	40	103	185	215	189	187	930	

FIGURE 2. COMPARISON OF OPPORTUNITIES WITH VOCATIONAL AMBITIONS


Percents
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70
H. S. pupils choosing professional careers
Workers in U. S. in professions
H. S. pupils choosing business and clerical positions
Workers in U. S. in business and clerical positions
H. S. pupils choosing mechanics and industry
Workers in U. S. in mechanics and industry
H. S. pupils choosing agriculture
Workers in U. S. in agriculture

cultural, mechanical and industrial occupations include 11 different employments, and 8.8 percent of the choices. See Figure 2.

Vocational opportunities, as shown by the United States census reports,[10] are just about the reverse of the distribution of high-school pupils' occupational choices. Agriculture and the mechanical and industrial arts engage the energies of 61.1 percent of the gainful workers in the United States, and only 8.8 percent of the high-school pupils had ambitions looking toward these fields. Business and clerical employments enlist only 14.1 percent of the gainful workers of the country; and yet 29.7 percent of the high-school pupils plan to enter these fields of effort. In the United States census reports practically all of the occupations designated in Table II as unclassified, are included under the caption "professional service." In spite of this liberal interpretation of the term "professional" only 4.4 percent of the gainful workers are found to be engaged in professional service in this country. Nevertheless the professional and unclassified divisions of Table II include 574 choices, or 61.7 percent of the total number.

Although it is a fact that the high school represents a rather highly selected group of young people from whose ranks the clerical, business, and professional occupations are very largely recruited, it is apparent that by no means all of the 91.2 percent of high-school pupils will find their way into these occupational fields. Furthermore, for their own best good and the best good of the nation a great many of them should be directed toward the agricultural, mechanical, and industrial fields.

The question may be raised as to the use that a vocational counselor might make of such facts regarding the intelligence of high-school pupils as are shown in Table II. Assuming that occupational levels approximating those found in Table I have been established, the counselor could proceed on the theory that those falling within the lowest quartile of intelligence ratings, i.e. who make scores lower than the beginning score for the middle 50 percent of a given occupation, would probably have small chances of success as workers in that occupation.

[10] *Thirteenth census of the United States, 1910. Volume 4: population: occupation statistics.* Washington: Government Printing Office, 1914, p. 40.

Take the professional group. The intelligence score limits of the middle 50 percent as shown in Table I are from 98 to 184 (Alpha). There are fifty cases in the professional group, Table II, who made an intelligence score of less than 90 points on the Alpha scale. The counselor could safely consider the cases falling in this quartile of the professional group as doubtful. If teachers' estimates of intelligence and school marks confirmed the findings of the mental tests, he would be justified in making an effort to direct the thoughts of those boys and girls toward some other life career.

Specific occupations treated in the manner just outlined for the entire professional group would yield the following percents of doubtful cases: draughtsmen, 23.8 percent; bookkeepers, 21 percent; stenographers 22.1 percent; dentists, 18.1 percent; engineers, 24.3 percent; lawyers, 15 percent; doctors, 13.9 percent; and teachers, 33.6 percent.

If by making use of intelligence ratings of individual pupils in connection with the intelligence levels of the occupations which they are ambitious to enter the vocational counselor can give them more accurate advice as to the life career in which they are most apt to succeed he should by all means make use of such ratings. But it will probably be pointed out that intelligence ratings are not in themselves sufficiently reliable to justify their use in vocational guidance. To this objection it may be replied that the high-school counselor has at hand means of verifying the results of the mental tests. He can secure the estimates of teachers and others who know the pupils and he can secure the record of their success in school tasks. Agreement between teachers' estimates of intelligence and mental tests, or agreement between school marks and mental tests would greatly strengthen the presumption that the tests had succeeded in discovering the pupil's mental level.

Discovery of Agreement between Mental Tests, School Marks, and Vocational Ambition

The cases of the high-school pupils, whose mental ratings and vocational ambitions are set forth in Table II, are redistributed in Table III in accordance with standing in school subjects, intelligence tests, and rank of vocational ambition. For the purposes of the three-way distribution, vocational ambitions

are ranked as follows: Rank I: higher professional and executive positions; Rank II: business, semi-professional, higher clerical positions; Rank III: general clerical, skilled labor etc.; Rank IV: semi-skilled labor; Rank V: unskilled labor.

Mental age equivalents for the Army Scale ratings were worked out by Dr. Samuel C. Kohs and the writer, and from these intelligence quotients (I. Q.'s) were computed. The mental level of a child is more nearly represented by the I. Q. than by the raw scores on an absolute point scale because the latter take no account of the chronological age.

Sex differences in rank of vocational ambition are indicated in column 10. There are 189 boys, or 45.1 percent of the total number of boys, but only 86 girls, or 16.8 percent of the girls, who chose vocations of the first rank. The vocational ambitions of 33.4 percent of the boys and 46.3 percent of the girls are represented by Rank II, and those of 21.5 percent of the boys and 36.9 percent of the girls by Rank III. There were none of the 930 who chose vocations of less than Rank III. The most numerous choices of the boys were for the engineering profession and of the girls for stenography and teaching.

How those whose ambitions come in the different ranks would probably measure up in mental ability to the demands of the occupations chosen can be estimated by reference to the nine possible combinations of I. Q. and school marks. Entries in column 1 show those who have I. Q.'s and school marks both below average. There are 51 classes in this group. Seven of them have chosen occupations of Rank I, and 18 of Rank II. Since school marks confirm indications of mental tests as to low mental level in these cases the success of these pupils in occupations of Ranks I or II would be open to question.

In column 8 there are 239 cases, 77 boys and 162 girls, having marks above average and I. Q.'s average. Here is an indication that the tests did not register the full ability of the pupils, or that they possess qualities of persistence and other attributes tending to supplement intelligence as factors in successful school work

Columns 2 and 3 indicate that there are many pupils whose intelligence is average or above average but who do poor school work. Teachers are apt to rate such pupils low in intelligence.

TABLE III. DISTRIBUTION OF 930 HIGH-SCHOOL PUPILS ACCORDING TO INTELLIGENCE, SCHOOL MARKS, SEX, AND VOCATIONAL CHOICE

Rank of Vocational Ambition	Marks Below Average						Marks Average						Marks Above Average								
	1 I. Q. Below Average		*2* I. Q. Average		*3* I. Q. Above Average		*4* I. Q. Below Average		*5* I. Q. Average		*6* I. Q. Above Average		*7* I. Q. Below Average		*8* I. Q. Average		*9* I. Q. Above Average		*10* Totals		
	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	All
I	7	0	13	4	17	3	7	4	18	3	15	3	4	8	32	25	76	36	189	86	275
II	7	11	29	18	4	7	4	14	10	34	16	4	5	14	27	84	38	51	140	237	377
III	6	20	14	26	5	0	3	11	15	30	13	4	7	19	18	53	9	25	90	188	278
IV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	20	31	56	48	26	10	14	29	43	67	44	11	16	41	77	162	123	112	419	511	930

The mental tests give the counselor an insight into their true ability and enable him to employ methods of waking them up through the discovery of the right life-career motive.

The cases entered in columns 6 or 9 where the I. Q.'s are above average and school success is average or above average can be considered as satisfactory, so far as intelligence is concerned, for the occupations chosen. However, the counselor can be of great service to the pupils in these groups through his ability to give information as to the demand for workers, the opportunities for advancement, the qualifications as to health, temperament, training, etc., expected of those who enter the occupations selected for consideration.

There is no purpose here to suggest that a counselor should always advise those who have superior ability to enter high ranking occupations. If such a boy or girl would be more contented in an occupation ranking low on the rating scale there is no occasion to urge him to select another simply because it ranks higher in popular estimation. There is room for superior intelligence in every occupation, and it would be well for the country if young people of superior ability were encouraged to follow agricultural, mechanical, and industrial pursuits, in order that they might become leaders therein. The professions are overcrowded, but there is always room for young people with ideas and energy in the food producing and industrially creative occupations.

Summary and Conclusions

1. The use of psychological and trade tests in the selection of employees can be more easily shown to be effective, than can the use of such tests in vocational guidance. The employment manager can "play safe" by rejecting all but the most promising applicants, while the vocational counselor must needs endeavor to give sound advice to all comers.

2. Minute charting of abilities by means of psychological and trade tests is not practicable at the present time for the public school vocational counselor. There are far too many different occupations and the specific abilities of individual pupils are much too various to permit of accurate "pigeonholing" according to manual, conceptual, and other types. Likewise a given combination of abilities might mean successful participation in any one of a wide range of occupations.

3. The discovery of the levels of intelligence of individuals and of occupational groups may prove to be of great assistance to the high-school counselor. The data on army mental tests, arranged in Table I, indicate that there are rather definitely marked occupational levels of intelligence. The norms already suggested would probably be confirmed by a compilation of all available data.

4. Application of the Army Intelligence Scale to 930 high-school pupils and the distribution of the cases according to intelligence rating and vocational ambitions is shown in Table II. Illustration is also given of the way in which this knowledge might be applied to the vocational guidance of the group tested.

5. Need of vocational guidance of high-school pupils is shown by the fact that the concentration of choices in the professional, business, and clerical occupations is out of all proportion to the opportunities in those lines as shown by the United States census. The demands for workers in agricultural, industrial, and mechanical pursuits should be emphasized by the counselor.

6. The need for supplementary information to verify the findings of the intelligence tests is shown in Table III. The 930 cases are distributed according to sex, school marks, vocational ambition, and intelligence. Where intelligence rating and school marks agree the presumption is that the intelligence level was approximated by the mental test. Where they do not agree it is a warning to make further inquiry into the matter.

7. The employment of psychological tests as an aid in vocational guidance is in the early experimental stage, but sufficient progress has been made to justify their use in a negative way, i.e., as a means of discovering to the counselor the kinds of occupations that a given high-school pupil would probably better avoid. They are useful also as a means of satisfying a counselor that a given pupil has the mental ability to engage in the occupation which he has chosen, providing other necessary factors conditioning success are present. In any case the counselor will do well to remember that he is dealing with probabilities and not with certainties. The mental tests, if conservatively employed, will increase the probability that the counselor will give really helpful advice.

STANDARDIZING PROCEDURE IN A COMBINATION SYSTEM OF SUPERVISED STUDY, VARYING SCOPE OF WORK, AND WEIGHTED CREDIT

W. H. HUGHES
Principal, Riverview Union High School, Antioch, California

The purpose of this paper is to set forth some of the methods employed in the Riverview Union High School in administering a combination plan which takes into account the variations of natural capacities and acquired interests, and at the same time accords credit toward graduation in proportion to individual achievement. The more general features of the plan are included in the accompanying chart taken from our *Teachers' and Students' Handbook*.

Provisions of the general scheme summarized.—The more outstanding provisions of the system may be briefly summarized as follows: (1) a specific statement in each subject concerning the scope and quality of work which a student of average ability should be able to accomplish when working up to his full capacity, i.e., the general requirements for the medium grade, B, which entitles the student to 1.0 unit of credit; (2) a careful enumeration in each subject of the characteristics of excellent work entitling the student to the A grade which carries with it 1.1 units credit; (3) a general statement in each subject relative to the many possibilities for extra scope of work which the student of superior ability, interest, initiative, and application may do along the lines of his special interests in satisfaction of the extra requirements for the AA grade which entitles him to 1.2 units credit; (4) special encouragement and direction for the student of extraordinary ability, interest, application, and achievement in doing an original piece of work relative to the subject for which he may be entitled to the AAA grade and 1.3 units credit; (5) provisions for diminished credit for inferior work; and (6) a full hour class period the first half of which is given to recitation and discussion and the last half to supervised study, enabling the teacher not only to give attention to the backward student in bringing him as nearly as

possible up to the medium standard but to direct and supervise the efforts of the more capable student along the lines of his special interests. It will be noted that the plan is not merely one for

EXPLANATION OF MARKING AND CREDITING SYSTEM
(From the *Teachers' and Students' Handbook*)

Marks	Meaning of These Marks and Suggestions for Improvement	Crediting Value for Two Semester Subject	Approximate Number of Students in 100 Receiving This Mark
AAA	The AAA mark is reserved for the student or students of extraordinary ability, interest, initiative, and application. In addition to meeting the requirements for all lower marks, the AAA student must do a piece of original work in which he shows his ability to collect information from a variety of sources and to organize it into a well-written paper of reasonable length. Or, if the subject is non-academic in nature, the extra piece of original work may differ from that suggested above but, in any case, must be approved by the head of the department and by the principal of the high school. One or possibly two students in one hundred may reasonably be expected to do the work required for this mark.	1.3	1 or 2
AA	The AA mark is reserved for students of superior ability, interest, initiative, and application. In addition to meeting the requirements for all lower marks in a given subject, the AA student must cover an extra scope of subject matter which is to be indicated by the teacher. This extra scope should be rich in practical applications of the subject and should lead the student into a systematic use of supplementary materials such as reference books, encyclopedias, periodicals, government reports, etc. The nature of this extra scope should, so far as possible, be determined by the special interests of the student. Approximately, four students in one hundred should be successful in doing the work required for this mark.	1.2	4
A	The A mark is used to denote excellency in meeting the general requirements of a given subject. In a group of one hundred students there should be approximately twenty receiving the A mark. This twenty should be better than the seventy-five who receive lower marks; but hardly so good as the five best students in the hundred. The student receiving the A mark should show ability, interest, self-reliance, and application superior to the average student. His recitations should be excellent; his written work, accurate, neat and legible; his attendance regular; and his attitude toward school commendable. Nothing in the explanation of this mark should be considered to preclude assignments of desirable supplementary work.	1.1	20

MARKING AND CREDITING SYSTEM *(Continued)*

Marks	Meaning of These Marks and Suggestions for Improvement	Crediting Value for Two Semester Subject	Approximate Number of Students in 100 Receiving This Mark
B	The B mark is used to denote average ability and attainment in a given subject. In a group of one hundred students approximately fifty may be expected to receive this mark. This means that there are twenty-five students better and twenty-five poorer than this medium group of fifty. To receive the B mark a student must satisfy the general requirements of the subject with respect to the specific assignments, oral and written work, exercises, note books, translations, book reviews, tests, etc. All work must be reasonably prompt, neat and accurate. Good attention in class and proper use of time in school hours are taken into consideration in the assignment of this mark.	1.0	50
C	The C mark is used to indicate ability and attainment somewhat below the medium. In a group of one hundred students, approximately seventy-five will do a quality of work better, and five will do a quality of work poorer, than that represented by the mark. The quality of work entitling the students to this mark is just as far below the average as the quality of work for the A mark is above the average. For the marks C and B the scope of work is practically the same. The C student wishing to improve the quality of his work should follow the suggestions given under A and B.	.9	20
D	The D mark is used to represent the next lower step in the quality of work accomplished. It is as much inferior to C as C is inferior to B. In a group of one hundred students approximately four may be expected to do work of this grade and character. The D student who wishes to improve the quality of his work should follow the suggestions given under B and A.	.8	4
E	The E mark is used to indicate the poorest quality of work which may be accepted for credit toward graduation. It means that the student is either inferior in his ability or that he is not making the proper use of his time. The E students should try to improve the quality of their work by following the suggestions given under A and B.	.7	1
F	The F mark is used to indicate absolute failure.	.0	?

giving "extra credit for extra quality" of work accomplished. Quality is important but is not in itself sufficient. *Extra scope* and

extraordinary kind of work are to be encouraged if the student of superior ability is to receive justice.

Standardizing teachers' procedure.—One of the first steps in the administration of such a system is to make teachers familiar with the fundamental principles of variation. Although, on every hand, they are confronted with the general facts of variation, they need to have brought to their attention many specific cases of exact measurement and distribution. Illustrative cases for this purpose may easily be found in various books of statistical method, in the studies of Thorndike on individual differences, and in many articles of educational and psychological journals. Incidentally, the results obtained from various intelligence tests in our own high school still further corroborate the theory that capacities are distributed remarkably in accordance with the symmetrical curve.

The next move on the part of the administrator is to make clear the approximate distribution of marks in the adopted system if it is to be at all scientific. "We would laugh at a government," says Professor Max Meyer, "requiring each citizen to pay an income tax, but failing to state what percentage of the income is to be paid. We should likewise laugh at a school administration telling its teachers to report some A grades, some B grades, some C grades and so on. Each teacher, in reporting, would then unwittingly grade, not his pupils so much, but rather himself, his own notion of what ought to be done. That is, one would report that according to his conviction a teacher ought to report only, say, six percent A grades; another would express his conviction that a teacher ought to report, say, twelve percent A grades; a third would grade himself as believing in eighteen percent A grades; and so on to the most liberally grading teacher. But the school administration does not want its teachers thus to grade themselves. It wants them to grade their pupils, so that one may tell who is the best kind of scholar, who a less good scholar, and who a very poor scholar."

We have therefore taken the precaution not only to work out rather definitely the specific requirements for the different marks in our system but to indicate approximately the percent of students who ordinarily should be expected to receive these marks respectively. The B grade in our system is taken as the starting

point. (See definition in chart given above.) It is what the student of medium ability should earn when working up to his full capacity in school time and spending half an hour daily on each subject outside of classrooms. Since the student carries four subjects regularly, the teacher is cautioned against requiring more for the B standard than can reasonably be accomplished in half an hour of home study. (It is generally understood that on the average two hours of home study are necessary for satisfying our general requirements.) According to our definition of the B mark approximately fifty students in one hundred (in addition to those securing A, AA, and AAA) should earn it.

The B standards having been indicated, it is equally important to make clear the general requirements for the higher marks. The teacher is occasionally reminded that the A mark is used to denote *excellency* in meeting the general requirements of a given subject. (See statement concerning this mark in the chart at beginning of this article.) According to our definition of this mark approximately two students should earn it for every five who receive the B mark. This number, however, may vary; but any marked variation receives the attention of the principal. The AA mark, it will be noted, stands for *more than mere excellency* in meeting the general requirements of a subject. It is reserved for students of superior ability, interest, initiative, application, and achievement *who master an extra scope of subject matter or work* in a given study. The extra scope must be approved by both principal and teacher. This provision serves as a check to teachers who might otherwise be inclined to assign too many AA marks. It helps also to standardize the amount and kind of work required for the mark. According to our definition of the AA mark, approximately one student should earn it for every five who receive the A mark. Any wide variation from this ratio requires explanation. In the same way, the AAA mark is reserved for the student or students of *extraordinary ability, interest, application, initiative, and achievement.* As stated above the student must offer a piece of individual work extending through the semester (or year) in addition to satisfying the requirements for all lower marks. *This extra piece of work must be approved by teacher and principal.* This method of approval serves as a check on the system and assures the maintenance of high standards. According

to our definition of the AAA mark approximately one or two students in a hundred should be able to secure it. All conditions being normal, a too large percent of the higher marks indicates the desirability of raising the standards.

Another important factor in the standardization of our marking system is a monthly comparison of the distribution curves of the different departments with the general curve adopted for our system. Our teachers know, however, that there are *certain acceptable reasons for departure from the theoretical curve:* (1) an unusually large percent of students in a given class may have come from cultured homes; (2) an unusually large percent may have come from uncultured homes; (3) students may be of unusual maturity of mind; (4) the teacher may be extraordinarily skillful in effective instruction; (5) the class may be inadequately prepared for the subject; (6) the class may be exceptionally well prepared for the subject; (7) there may be an unusual percent of superior or of inferior students. On the other hand, teachers know equally well that there are *unacceptable reasons for departure from the theoretical curve of distribution:* (1) low marks because of attempting to cover too much ground; (2) high marks because of meager requirements; (3) low marks based on inadequate data; (4) high marks to give an impression of efficiency; (5) high marks to curry favor with students and influential parents; (6) high marks to avoid criticism of troublesome parents; (7) low marks because of one's own disgruntled disposition; (8) high marks because of the student's "effort" although it has been unproductive. With so many possibilities for variation from a normal curve of distribution, it is evident that a system of weighted credits needs careful supervision in order to insure the maintenance of high standards.

When conspicuous variations from the adopted curve of distribution are in evidence it is, of course, the business of the administrator to ascertain if possible whether such variations are justifiable or remediable. In this connection we have found the results of intelligence testing of considerable value. It is needless to say that if a student has been found far below the average in mental capacity and is carrying away the highest marks in a particular class or classes, his case is one suitable for investigation. In like manner, if the student of superior ability as indicated by the

mental tests, is receiving marks considerably below the average, his case also needs attention. It is not always and necessarily found, however, that the marks of this bright student are too low. But the fact that there is a discrepancy between his capacities and his school standing may direct the attention of teachers and principal in such a way as to improve his work and eliminate the discrepancy. At the present stage of imperfection in mental testing, however, the principal should not depend exclusively upon the results of a single type of mental test. An average high score or an average low score for a number of the best-known mental tests is more dependable than a high score or a low score in any one of them.

As a mechanical means of administering the marking and crediting system, the accompanying form of report card has been devised and found useful.[1] The red lines are a constant reminder that B is a medium grade. Such a card makes it easy for the principal or supervisor quickly to get a general impression as to the comparative levels of grading of the different departments. With this card containing in graphical form the complete record of a student in all subjects for any month, semester, or full year, the principal can easily note any tendency of teachers to vary from the adopted standards. The key to the marking system included in every report card is also a checking device.

MONTHLY REPORT

To Parent and Student:

The purpose of this report is to indicate the relative standing of the student, with respect to the poorest and the best. The student's grade is indicated by a check [✓] mark under the appropriate letter. Any check under B between a pair of red lines means good, average work and achievement; a check to the right of the red lines means poorer than the average for students in general; a check to the left of the red lines means better than the average. In order to do creditable work, every student will find it necessary to spend at least two hours a day in study outside the class rooms.

W. H. HUGHES, *Principal*

[1] The card actually used carries an abbreviation of the "Explanation of Marking and Crediting System" given on page 548. In order to draw attention to the "B" rating as the average mark, the lines inclosing this letter are printed in red. The card is in the form of a four-page folder. The second and third pages are shown here. The first and fourth pages are reserved for the name of the school and of the pupil and for the parent's signatures.—EDITOR.

MONTHLY REPORT

	SUBJECT							
Mark	AAA	AA	A	B	C	D	E	F
Sept.								
Oct.								
Nov.								
Dec.								
Jan. First Semester								
Feb.								
Mar.								
Apr.								
May								
June Second Semester								
Final Standing								

	SUBJECT							
Mark	AAA	AA	A	B	C	D	E	F
Sept.								
Oct.								
Nov.								
Dec.								
Jan. First Semester								
Feb.								
Mar.								
Apr.								
May								
June Second Semester								
Final Standing								

The effectiveness of the foregoing methods may be seen in the following distribution of marks for the first semester of the school year 1919-1920. The entries are in percents.

Months	AAA	AA	A	B	C	D	E	F
September	0	0.5	10	49	27.5	9	2.9	1
October	0	4.0	16	48	26	5	1.0	0
November	0	3.9	20.1	44.3	24.7	5.6	0.6	0.4
December	0	4.1	19.7	43.5	23.0	8.3	0.8	0.4
January	0	2.0	17.8	42.2	26.0	9.5	1.8	0.5
Semester	0	3.9	19.5	43.3	24.8	6.6	1.8	0.0

With respect to supervised study, which is a part of our combination scheme, it is more difficult to standardize procedure because of the varied nature of the different subjects and the individual methods of teachers. Occasionally, however, we have found it profitable to call the attention of teachers to various possible sorts of classroom activities. Recently our teachers examined their own methods for an entire week indicating on a chart prepared for the purpose the approximate percent of class time devoted to each of the following activities: (1) assignment of work; (2) instruction in best methods of study, either general or specific; (3) recitation taking the form of quiz and consisting principally of questions and answers based on the general requirements in the regular text; (4) special reports by students on supplementary work related to the regular assignments, i.e., making the classroom a "clearing house" for information; (5) cooperative development lesson during which all students are simultaneously engaged under the leadership of the teacher in working up a new portion of the subject; (6) individual work by students either with supplementary books and materials or with laboratory and shop apparatus; (7) supervised study of general assignments and direction of extra-scope activities; (8) lecture by instructor; and (9) any other class activity employed by the teacher. The results of this study showed that our teachers are spending not more than one-fourth of the class hour in recitation, but at least three-fourths of the hour in the other forms of activity.

The attitude of our faculty and student body toward the system is most favorable. Almost without exception, our

teachers are enthusiastic over the plan and its possibilities. Many of us believe that the students of superior capacities are now accomplishing at least twice as much as formerly. We are equally confident that the work of every type of student has been improved. Never before have our teachers seen students so eager for supplementary work. Under the old system few of us believed that a freshman would, of his own accord, read over two thousand pages of supplementary material in a semester subject. Several freshmen under our present system have done this and at least one has read nearly three thousand pages. In every subject, students are "reaching out after more." These are some of the many reasons why practically all of our teachers are advocates of the present combination system of supervised study, varying scope of work, and crediting in proportion to individual achievement.

MEASUREMENTS IN FAIRY TALE[1]

CYRUS D. MEAD
University of California

Once upon a time there was in old England a very tall king who was about to make a fair princess the queen of his realm. Trousseaus were of as much interest then as now, so the Very Tall King and the Fair Princess went shopping down through the narrow streets of London Town. By and by they came to the Shop of the Medium Sized Shop Keeper.

"I want," said the Very Tall King, "twenty yards of gold brocade." The Medium Sized Shop Keeper immediately drew forth a large roll and began to measure out the twenty yards of cloth in the accustomed manner.

"Hold on there," said the Very Tall King, "don't you know that it is from the point of the *King's* nose to the tip of his fingers that a yard is measured?"

"That is true, your Majesty," said the Medium Sized Shop Keeper, and handing the roll across the counter, he watched ruefully while the Very Tall King measured out twenty great, long yards of gold brocade.

By and by, as time went on, the Very Tall King died and his very short brother came to the throne. He, too, sought a fair princess for his queen, and one day they visited the same Shop of the Medium Sized Shopkeeper.

"I want," said the Very Short King, "ten yards of your whitest satin." The Medium Sized Shop Keeper threw the bolt down upon the counter, folded his arms, and waited.

"I want ten yards," again said the Very Short King.

"Very well, O King, measure it out," returned the Medium Sized Shop Keeper. "A yard, you know, is from the point of the King's nose to the tip of his fingers."

"You are right," replied the Very Short King, and he measured out ten little, short yards of satin.

Such was the yard in merry old England and even today some people have seen the department store clerk measure goods in

[1] Being the introduction to a demonstration of tests before the San Bernardino City Institute, December 16, 1919.

lengths of arm reach from the nose to the finger tips. So yards may come and yards may go, as clerks may be very short or very tall, but is a yard a yard?

A few days ago I stood before an inclosed glass case in the chemistry building of the University of California. My attention was called to a steel-and-brass rod. It was the shape of a miniature railroad "T" rail. A card read: "This is one meter—one ten-millionth of the distance from the earth's equator to the pole." And so meters may come and meters may go, but a meter is a meter regardless of kings and clerks. Yes, a yard is a yard today but it was centuries in the making.

The physical sciences have been made exact and are sciences chiefly in that there has been evolved, for each, definite units of measurement.

At one time, however, man expressed quantity in such terms as "heap much"; but "heap muches" varied. Yesterday the city across the bay voted upon a municipal tax rate of three dollars and eight cents. It may be a question whether or not the city's money will be wisely spent, but each citizen at least knows just how much he is to contribute. Regardless of municipal incompetency and rottenness, we are wont to call this "civilization." However that may be, it took man ten thousand years, or more, in economics to go from "heap much" to three dollars and eight cents.

A long time ago man expressed time by such terms as "many moons." Yesterday some of the boys landed at the Presidio with a flying time from New York of forty-three hours, thirty minutes and fifty seconds. It took man ten thousand years, or more, in chronology to be able to change "many moons" to hours, minutes, and seconds.

At one time number was expressed in such terms as "like the leaves of the trees." Yesterday the registrar said there were 9,457 "campus" students enrolled in the University of California, making it the largest university in the country. The science of definite numerical calculation has changed "the leaves of the trees" to units, tens, hundreds, and thousands, and definite comparisons become possible.

Now we as teachers have long since dealt in yards and meters, dollars and cents, hours and minutes, tens and hundreds; but we still make use of very changeable and indefinite terms. We claim

that a pupil's work is poor, fair, good, or excellent; but what is poor, fair, good, and excellent? Why, you say, poor is 60 percent, fair is 70, good is 80, and excellent is 90. But what is 60 percent? It is not only one thing to one and another thing to another, but it is also different for each of us at different times. Grading as tangible a product as a piece of writing without a handwriting scale and then with one will prove this. One geometry paper marked by a hundred and sixteen teachers of geometry was scored all the way from 28 to 92 percent, and this in a day when the physics laboratory of University Chicago "can rule a piece of polished glass one inch wide with fifty thousand straight and parallel lines equally spaced" and "can measure the five-millionth part of an inch—a distance amounting to one fiftieth of the smallest distance revealed by a theoretically perfect microscope."

For many centuries man has dealt in relatively exact terms in many lines of endeavor, and in so doing he has lifted himself from the stone age to a civilized state. It may have taken him ten thousand years. We in education cannot—probably never can—deal in terms anywhere nearly as exact as those of the physical sciences; but the time we have spent in developing more precise units has only been since 1910 when Thorndike evolved the first scale for handwriting! It is now my purpose to demonstrate the uses of some such scales and standardized tests. Will the class please come forward?

PROVISIONS FOR INDIVIDUAL DIFFERENCES IN THE TEACHING OF READING[1]

W. W. THEISEN

Director of Educational Measurements, State Department of Public Instruction, Madison, Wisconsin

The results of standardized tests have everywhere revealed wide differences in reading ability. They have shown decided variations in such factors as rate of reading, knowledge of vocabulary, ability to gather thought from the printed page, and ability to read orally. Diagnostic analyses of results such as those made by Miss Zirbes[2] and others throw additional light on the nature of these differences. But in spite of all the facts brought to light, few attempts have been made in classroom teaching to provide adequately for individual differences in reading ability. The question is what can be done and what is being done by *successful* teachers.

The first question which this study seeks to answer is: What provisions for individual differences are employed by teachers generally considered successful in the teaching of reading? It is a matter of common observation that the average teacher does not satisfactorily make such provisions. For that reason we have tried to limit our study to strong teachers. The main purpose has been to find out what provisions are in actual use by such teachers and which provisions can be recommended to teachers in general.

Our second question is: What suggestions for the less experienced teacher can be gained from the study? E.g., What values have the various provisions which we have found to be in use? Are the most valuable always among those most favored? Which ones are suitable for the different grades?

Light was also sought on these additional questions: Are even these teachers of acknowledged success widely conscious of the problem? How far do they avail themselves of the means at hand?

[1] An address delivered before the National Association of Directors of Educational Research at Cleveland, Ohio, February 26, 1920.

[2] Zirbes, Laura. "Diagnostic measurement as a basis for procedure," *Elementary School Journal*, 18:505–22, March, 1918.

Method of the Study

A letter was sent to a number of supervisors and superintendents of schools where there was reason to expect good instruction in reading with the request that a few of the best teachers be asked to reply. Copies of a circular stating the nature of the inquiry were inclosed for the benefit of the teachers. The circular merely asked that the person write us in detail telling what means she was using in the teaching of reading to provide for individual differences. She was requested also to offer suggestions as to methods that she thought could be used. Of the 123 replies received 118 were used in the study. A few appeared to have misinterpreted our request. The 118 were distributed on a percentage basis as indicated in Table I.

The results of this inquiry will be referred to throughout as the *original study*. After the first group of returns had been tabulated, a suggestive list containing the provisions listed in Table I, without figures, was sent to those who had replied and to a number of other strong teachers. On this list each person was asked to check the provisions which he or she used in teaching. The object was to provide a check on the results obtained in the original study.

The Provisions Found

For the present we shall confine the treatment to the original inquiry. The provisions found may be divided for convenience into two general divisions. The first we may designate as *extra-classroom* and the second as *intra-classroom*. Those belonging to the first have to do more with the steps preliminary to actual classroom procedure. In the *extra-classroom* division are two groups of provisions that offer us some assistance in solving the problem of individual needs. They apply not only to reading but to all subjects. (1) Attention to physical needs. (2) A flexible system of promotions. The *intra-classroom* provisions may be classed under seven headings in Table I: (A) Testing of individual abilities; (B) Grouping of pupils within the classroom; (C) Gradation of materials; (D) Variation in amount of reading practice; (E) Personal attention given to individuals; (F) Adaptation to and development of individual interests; (G) Specific forms and phases of instruction—or what may be called "methods."

TABLE I. PROVISIONS FOR INDIVIDUAL DIFFERENCES IN THE TEACHING OF READING REPORTED ORIGINAL STUDY

Provisions	Percent Reporting the Use of Each Provision				
	Primary	Intermediate	Grammar	Supervisors	All
(A) *Testing of Individual Abilities*					
1. Tests of rate and comprehension	6	32	29	46	24
2. Intelligence tests	6	0	0	0	3
(B) *Grouping of pupils*					
1. Pupils arranged in small groups by ability	58	39	14	54	45
2. Pupils in small groups *not* by ability	3	7	14	8	6
(C) *Gradation of materials read*					
1. Graded to fit individual ability	42	32	7	46	31
2. Selected by pupil's own choosing	10	20	21	23	16
3. Portions in oral reading or dramatizing assigned according to individual ability	8	7	14	15	9
(D) *Variation in amount of reading practice*					
1. Voluntary reading in free periods	40	22	21	23	29
2. Home reading	27	44	29	23	32
3. Additional periods	15	5	7	23	11
4. Poorer called on more frequently in oral recitation	0	0	7	8	3
5. Rapid cover more in silent reading	4	7	7	0	5
(E) *Personal attention to individuals*					
1. Special help for specific defects	46	44	43	46	44
2. Use of better pupils to help poorer	38	20	14	23	26
3. Home help	4	2	21	0	5
(F) *Adaptation to and development of individual interests*					
1. Materials chosen for individual interests	8	22	29	31	18
2. Individual choice of parts	4	2	7	8	4
3. Best read for inspiration to poorer	2	5	0	0	3
(G) *Specific forms and phases of instruction*					
1. Centering attention upon thought					
a. Silent reading	65	83	71	77	65
b. Use of thought questions	27	51	43	15	36
c. Illustrate material read by drawing	6	7	0	0	5
d. Thought drills	3	27	21	8	14
2. Problem assignment	4	2	0	8	3
3. Motivation	54	54	36	23	48
4. Dramatization for certain pupils not included under C-3	3	10	7	15	7
5. Word and phonic drills	35	39	29	54	41
6. Phrase drills	19	29	7	31	23
7. Speed drills	0	17	29	23	14

Note: Percentages are computed on the basis of returns from 118 teachers and supervisors distributed as follows: Primary grades I and II, 48; intermediate, grades III, IV, and V, 41; grammar, grades VI, VII, VIII, 14; supervisors, 13; ungraded, 1; not specified, 1. Geographically they represent 13 states.

We shall dwell only very briefly on the two groups under the *extra-classroom* division, as these are usually outside of the control of the classroom teacher. As regards health, it may be said that cases of individual differences often call for physical attention. Every school system should provide for expert diagnosis and treatment of physical defects. Many teachers believe that pupils would receive instruction more nearly suited to their needs under a genuinely flexible system of promotion. This would undoubtedly aid in the solution of our problem. But these provisions are usually administrative matters, and as such are outside of the scope of our study.

Let us now consider the measures which we have called *intra-classroom* or teaching. The various provisions under the seven group headings are given in Table I. The percent of frequency for the different classes of teachers and for all combined is shown. *Testing of individual abilities* is placed first because it is logically the first step. Adequate provision for individual differences cannot be made without testing and analytical study of the abilities of pupils. Diagnosis, in education as in medicine, must precede prescription. It is significant that only one-fourth of the teachers considered the application of educational tests of sufficient importance to be mentioned. The use of intelligence tests was indicated by but 3 percent of them. We may be charitable in the case of intelligence tests as few teachers are either trained to give them or able to call upon anyone else to do so. But we can no longer excuse any but possibly first-grade teachers for failure to apply educational tests and to utilize the findings as a basis for intelligent procedure.

We have as one proposal for meeting individual differences *the grouping of pupils of a grade into several small groups*. Grouping together pupils of approximately equal abilities was among the three provisions most frequently mentioned. Another form of grouping is sometimes used, although it is of somewhat doubtful merit as a specific provision for individual differences. It is used in oral reading recitations, and consists in dividing the class into smaller units without attempting to place pupils of equal abilities together. The underlying theory apparently is that individual needs are met in part if the individual's opportunity for reading is increased.

The particular merit of grouping on the basis of ability is that it permits *gradation of materials*. If pupils having similar weaknesses and instructional needs are grouped, the treatment of their special deficiencies can be carried on apart from the remainder of the class. An application of this method may be seen in the study by Miss Zirbes to which reference has already been made.

Gradation of materials furnishes an important means of providing for differences in ability. The aim here is to allow each pupil to read material suited to his development. Thirty-four percent of the teachers reported that they attempt to choose material fitted to individual abilities. Nine percent say that they attempt to base the assignment of portions of the lesson in oral reading or in dramatizing upon individual abilities. Sixteen percent allow pupils to select materials for themselves. Naturally a child rarely chooses material so difficult that he does not enjoy it.

Individual differences in reading ability can be cared for in part also through *variations in the amount of reading practice*. Variations in practice can be afforded through voluntary reading in free periods, home reading, group assignments, and individual assignments. Some provide additional class periods. The brighter may read with an advanced section. The weaker may read with a lower section or remain after school for an extra period. The advisability of the latter may be questioned. Calling upon the poorer pupils more frequently in the oral recitation, as employed by a few of our correspondents, is indefensible from the standpoint of the brighter pupils.

Personal attention is a common method of meeting the requirements of individuals. By this means, it is hoped to give each child or group such attention as is needed to remove specific weaknesses. Mass drill is apt to be uneconomical to say the least. Pupils can be assisted with their difficulties through individual or group attention on the part of the teacher (while others engage in silent reading), through assistance from stronger pupils, and through home help. Of the three types listed in the table special help from the teacher is most often given. The use of the better pupils to aid the poorer, a practice that is open to question, is mentioned all too frequently. Too many teachers seem obsessed with the notion that abilities should be "levelled off" as one teacher expressed it, rather than permitted to make a natural

development. Home help does not appear to be popular with teachers.

In the judgment of a number of our correspondents, individual reading needs can be satisfied further by *developing and utilizing the child's own interests* and choosing his reading material accordingly. The table shows, however, that only a small percent suggested provisions that could be classed under interest. Only 18 percent, ostensibly, are making efforts to have material chosen in accordance with individual interests. The showing of the primary teachers is lamentable. The teacher must seek to discover a child's impelling interests as a basis for the choice of materials for him. In the words of one teacher, "Every child shows more interest in some one thing than another. The greatest difficulty is experienced in trying to find where this natural interest lies. This is the key which must be found in order to help him interpret new experiences."

To meet individual interests adequately every school must provide an abundance of material. This is where many fail. Good material for primary grades remains largely to be written. It is almost impossible to find any considerable number of books based upon worth-while interests of children and suitable for first or second grades. *We impose upon the good nature and submissiveness of children by putting them through the dry, artificial, and disjointed material of method readers, or fool them into liking stories founded upon superficial interests and built upon unreal situations.*

The final group of provisions of Table I entitled *specific forms and phases of instruction* includes provisions that pertain more to the teaching process. Our aim should be to furnish such teaching as will most nearly fit individual needs and permit maximum individual development. This can be accomplished, to an extent, by choosing methods that permit independent progress. This is doubtless the thought underlying "silent reading," which was suggested more often than any other single provision. The problem method is especially suitable for permitting independent progress. The unfamiliarity of teachers with this method and with its possibilities is emphasized by the fact that only 3 percent of the teachers mentioned it.

Our program for individual development can be carried out further by the choice of methods that center attention primarily

upon thought and the development of right habits of thinking. The child who develops such habits becomes more nearly independent. These considerations underlie such provisions as are listed under the caption of "centering attention upon thought."

We can do something for children by supplying incentives that lead to independent efforts. Motivation in some form is quite prominent with each class of teachers represented in the table. Some of the suggestions classed under this heading include the preparation of a story to be read to the class, test charts upon which the individual can note his own progress, dramatizing (especially in lower grades), and contests of various sorts.

Instruction for the pupils who need it in particular phases of reading, such as word analysis, word differentiation, speed, phrasing, organizing of the thought, and the like tends toward a proper individual development. Some of the provisions of this type could possibly have been classified under the caption of "personal attention." Word and phonic drills for individuals rank relatively high in the whole list of provisions.

A better idea of the frequency with which provisions in each of the seven groups were mentioned by each class of teachers may be had from the summary in Table II.

TABLE II. PERCENT OF TEACHERS WHO STATE THAT THEY MAKE SOME PROVISION CLASSIFIED UNDER THE GROUP HEADINGS INDICATED. ORIGINAL STUDY

Provision for:	Primary	Intermediate	Grammar	Supervisors	All
(A) Testing	10	32	29	46	25
(B) Grouping	60	44	21	62	49
(C) Gradation of material	44	54	36	54	47
(D) Practice	46	54	43	54	49
(E) Special help	60	61	57	62	59
(F) Interest	15	27	36	31	24
(G) Method	90	98	93	92	93

The greatest possible, and in most cases the most desirable, figure would be 100 percent. More than 90 percent suggested some provision which can be classed under *method* but only half, roughly speaking, suggest any form of provision that may be

classed under the headings of *grouping*, *gradation of material*, *practice*, or *special help*. The last is next to *method* in frequency. Only one-fourth list anything under the heading of *testing* or *interest*. The figures of the table, we fully believe, may be taken as indicative of what these teachers actually do. No large percent of them appear to have anything like an adequate comprehension of the problem. If these figures represent the actual facts for teachers of more or less acknowledged success, what must be true of *average* teachers!

The small percent of those who suggest *testing* is probably due in part to the fact that testing is rather new. This is especially evident in the case of the primary teachers. We are very much surprised, however, that fewer than a third of the intermediate and upper grade teachers thought of testing as a step toward adequate provision for individual differences. It is positively known that some in whose rooms reading tests had been given within the year did not mention testing. Exposure does not necessarily affect practice. *The problem of research workers is much more than inducing people to administer tests. They must be taught how to apply the findings.*

Grouping is much more popular with primary teachers and supervisors than with other classes of teachers. *Gradation* of materials drops off in the grammar grades where it should be strong.

The element of interest in primary grades should be a cause for concern. Only 15 percent apparently make serious efforts to fathom the interests of the individual child. It is entirely possible that this reflects the unusual deference paid to *mechanical methods* which appear to be blotting out all conceptions of the importance of interest. In the grammar grades, where teachers are forced by circumstances to study interests, the percent opposite *interest* is more than twice what it is in the primary grades.

By way of comparison it will be interesting to note what happened when these teachers were furnished with a list of all of the significant provisions already found and were asked to check the ones they were using. Table III gives facts similar to those of Table II, but for the suggested list. It is evident from this that teachers either failed on the original inquiry to record all that they were doing, or that they were ready to acquiesce very largely

TABLE III. PERCENT OF TEACHERS WHO STATE THAT THEY MAKE SOME PROVISIONS UNDER THE GROUP HEADINGS INDICATED. SUGGESTED LIST

Provisions for:	Primary	Intermediate	Grammar	Supervisors	All
(A) Testing	35	81	100	91	74
(B) Grouping	94	76	100	100	89
(C) Gradation of material	94	100	100	100	98
(D) Practice	100	100	100	100	100
(E) Special help	100	95	82	100	95
(F) Interest	100	86	100	100	93
(G) Method	100	100	100	100	100

in the privisions suggested to them. It is our candid opinion that Table II represents more nearly what is actual practice. It indicates what these teachers have thought about sufficiently to put into practice. Table III represents more nearly an ideal.

A further check on the breadth of teachers' conceptions of the problem may be had by noting the number of groups of provisions to which they contributed. Table IV indicates for both the original study and the suggested list the percents of each class of teachers who suggested provisions: (a) under four or more of the seven groups used in preceding tables; and (b) under three or fewer of them.

TABLE IV. PERCENT OF TEACHERS WHO CONTRIBUTED TO VARIOUS NUMBERS OF THE SEVEN GROUPS LISTED IN PRECEDING TABLES

Type of Teacher	Number of Groups of Provisions			
	Original Study		Suggested List	
	3 or less	4 to 7	3 or less	4 to 7
Primary	54	46	0	100
Intermediate	42	59	5	95
Grammar	64	36	0	100
Supervisors	23	77	2	98
All	48	52	2	98

In the original study nearly half failed to qualify for entry in more than three groups. The grammar-grade teachers appear to have the most limited viewpoint, for nearly two-thirds of them

failed to offer anything which could be included in four or more groups. The supervisors and superintendents give evidence of their broader conception of the problem. The summary for the suggested list is only what we should expect after seeing Table III.

When the teachers checked the suggested list, they gave us a cue to the measures which, in general, are most frequently favored. Table V indicates, in the order of frequency, the percent of teachers who checked each of the 28 provisions. Those most often checked by teachers of different types are indicated by a plus sign and those most seldom by a minus sign. There were but 61 returns, however, and for that reason the results must be discounted. Moreover, the wording of certain of the provisions may have caused some misinterpretations. The fact that we were dealing with provisions for individuals and not with class provisions may not always have been kept clearly in mind.

As any keen student will detect, the order falls short of being one of merit. We can judge from it which of the measures are relatively popular with each class of teachers and which ones are either not favored or are not sufficiently familiar to them. The five provisions most favored by all teachers are *silent reading*, *use of thought questions*, *voluntary reading in free periods*, *home reading*, and *special help for specific defects*. Strangely enough, voluntary reading appeared to occupy a much lower position among upper-grade teachers. It is possible that this is due in a measure to the formal programs of these grades, which permit little freedom or initiative. Among primary teachers *voluntary reading* receives first place. *Silent reading*, *special help*, and *grouping by ability* come next. Intermediate-grade teachers place *silent reading* and *use of thought questions* first. *Voluntary reading* is third. Upper-grade teachers place *educational testing* and *home reading* first. *Silent reading*, *use of thought questions*, *thought drills*, and having *the best read for inspiration to the poorer* are other provisions that rank high with these teachers. Every supervisor and superintendent checked *silent reading*, *thought questions*, *special help*, and *grouping by ability*. As might have been surmised, they checked a larger number of items than any other group.

The items checked by fewer than one-half of the 61 teachers will be seen from the table. It is to be regretted that so few

TABLE V. RANK OF THE VARIOUS PROVISIONS FOR INDIVIDUAL DIFFERENCES—BY PERCENT OF FREQUENCY. SUGGESTED LIST

Provisions	All	Primary	Intermediate	Grammar	Supervisors
*Silent reading (G)	97	94(+)	100(+)	91(+)	100(+)
Use of thought questions (G)	93	82	100(+)	91(+)	100(+)
*Voluntary reading in free periods (D)	87	100(+)	95(+)	45	91
*Home reading (D)	87	71	90(+)	100(+)	91
*Special help for specific defects (E)	87	94(+)	81	73	100(+)
Oral reading or dramatizing assigned according to individual ability (C)	85	77	90(+)	82	91
Best read for inspiration to poorer (F)	84	88	76	91(+)	91
*Pupils in small groups by ability (B)	82	94(+)	62	82	100(+)
Poorer called on more frequently in oral recitation (D)	79	71	81	73	91
Rapid cover in silent reading (D)	74	77	71	73	82
*Material graded to individual ability (C)	74	77	62	82	82
*Thought drills (G)	72	59	71	91(+)	82
*Tests of rate and comprehension (A)	70	29(−)	76	100(+)	91
*Materials chosen according to individual interests (F)	69	65	67	73	82
Material of pupils' own choosing(C)	69	65	71	45	91
Phrase drills (G)	69	82	71	36(−)	73
Word and phonic drills (G)	69	88	62	27(−)	91
Illustrate material read by drawing (G)	66	77	57	45	82
Use of better pupils to help poorer (E)	62	59	71	45	73
Speed drills (G)	61	65	67	45	64(−)
Motivation (G)	61	65	57	45	82
Individual choice of parts (F)	61	59	52	73	73
*Problem assignment (G)	48	29(−)	48(−)	64	64(−)
Home help (E)	48	47	43(−)	27(−)	82
*Intelligence tests (A)	48	18(−)	52	73	55(−)
Pupils arranged in small groups *not* by ability (B)	46	29(−)	43(−)	36(−)	91
Additional periods (D)	34	47	14(−)	9(−)	73
Dramatization for certain pupils (G)	33	29(−)	29(−)	45	36(−)

NOTE: The provisions starred represent those which we believe to be among those most worthy of consideration. Capital letters following each provision indicate the group under which it was classified—(A) testing; (B) grouping of pupils; (C) graduation of materials; (D) amount of reading practice; (E) attention given to individuals; (F) development of individual interests; (G) instruction. Figures followed by plus signs indicate provisions most frequently favored by each class of teachers. Similarly, minus signs indicate those least favored.

The figures above are to be taken merely as suggestive. They are based upon only 61 returns as follows: primary 17; intermediate 21; grammar 11; supervisor 11; mixed 1.

teachers are familiar enough with the *problem method* to put it into reading practice. This method could be used more successfully than most of those contained in our list as a means of providing for individual differences. The application of *intelligence tests* as we might expect occupies a low place because teachers have not been trained to use them.

The provisions which the writer deems among those possessing on the whole the most merit have been starred in the table. Several others are good. Few should be discarded entirely. These conclusions are based not only upon the study but upon observations of the classroom efforts of a number of teachers who were consciously attempting to meet the problem.

Our main purpose in this study, however, has not been to evaluate the provisions in use but to find, if possible, those that are being employed by successful teachers together with the suggestions to be gained for the less-experienced teacher. We believe the measures for meeting individual differences listed in Tables I and V represent a variety that will be suggestive. It is not a complete list. The validity of several of the proposed measures needs to be tested experimentally. Our best hope is that the study may serve as a stimulus to further investigation in this field. We trust that it will cause a few teachers at least to form a broader conception of the problem of meeting individual differences in reading ability.

SCALE OF ATTAINMENT NO. I.— AN EXAMINATION OF ACHIEVEMENT IN THE SECOND GRADE[1]

LUELLA W. PRESSEY
Indiana University

I. THE NEED FOR SUCH AN EXAMINATION

Systematic surveys of schools or school systems, in which standardized tests in the various subjects are employed, now constitute a common method of appraising the "efficiency" of a school, or system, in its work of training children. Relatively little has thus far been done, however, by way of putting together in convenient form groups of tests that may be thus used. In planning for a survey a superintendent ordinarily orders a variety of tests from a variety of sources—and gives them, often, in varying order. The directions from test to test are not related, sizes of blanks often differ, and, in order to gather together the results from such a survey, several separate blanks must be handled for each child.

The purpose in developing the scale to be described in the present paper was, first of all, to remedy these practical difficulties. School surveys have already proved their value. The development of instruments which would make such surveys less laborious and more practicable, and which would make the data obtained more readily interpretable with reference to every-day school problems, would surely seem well worth while. The writer has had certain further possible uses in mind in the development of the scale—uses which will be discussed shortly. But the examination is intended primarily as an instrument for use in the second grade in school surveys.

Effort was concentrated on the second grade because there is a lack of satisfactory tests for this grade, because the very concrete nature of the subject matter in this grade made test formulation relatively easy, and because the grade was considered especially important as that in which the first serious attack is made upon the "tool subjects." The grade also seemed especially appro-

[1] Studies from the Psychological Laboratory of Indiana University.

priate for a first experimentation as to the value of the general concepts involved in the examination. In the second, rather more readily than in any other grade, it should be possible to express in a single examination the fundamental features of the basic subjects taught.[2]

II. The Scale

The scale consists of four tests, one on each page of a four-page folder. On the first page are the spaces for name, age, grade, and sex, and then 24 lines on which are to be written the 24 words of the spelling test. On the second page is a test designed to measure reading vocabulary. It consists of 29 lines, 5 for a preliminary exercise and 24 for the test proper. Each line consists of four nonsense combinations and one real word. The children are told to "find the real word, and draw a line around it—so I'll know where it is." The first five lines of the test are reproduced below:

II. Word Recognition

1. pl bn is ct lu
2. ot the bis ln tsp
3. I e t m b
4. ml sb tu ro go
5. mp mbs ris see elo

The third page presents the arithmetic test, which consists of forty-six addition and thirty subtraction combinations. The fourth page gives the "silent reading" test, which is made up of 24 sentences (plus 5 preliminary examples), each one having an extra or confusion word in it; the children are to find this unnecessary word, and cross it out. The first five lines of the test run as follows:

IV. Understanding of Sentences

1. I ball see a boy.
2. See dog my pretty book.
3. I like to gun play ball.
4. I cat can see my doll.
5. The baby is my book little sister.

[2] It should be mentioned, however, that effort in this direction has not been limited to the second grade; the Laboratory now has in process of standardization an examination in American history, arithmetic, geography and English for use in junior high school, and plans further work along this line later.

Directions for the spelling and arithmetic tests are obvious enough and need no comment. The directions for the tests of reading vocabulary and understanding of sentences employ the question-and-answer method in developing the problem of the test, using the first five lines as examples. The directions for the reading vocabulary test are as follows:

Look at the first line! Who can see a real word in that line? What is the word? . . . Yes, "is." What part of the line is it in? . . . Yes, right in the middle. I want you to draw a line around the word "is" because it is the only *real* word in the line. The other groups of letters don't make words. Draw a line around "is."

Look at the next line. What is the real word? . . . Yes, "the." Draw a line around "the" because it is the only real word in the line.

Look at the next line. What is the real word? . . . Yes, "I." Draw a line around "I."

Look at the next line. What is the real word? . . . Yes, "go." Where is the word "go." Yes, on the very end of the line. Draw a line around "go."

Look at the next line. What is the real word? . . . Yes, "see." Draw a line around "see."

Now, everyone attention! I want you to read over the rest of the lines on the page in the same way. In each line there is *just one* real word. You find the real word in each line, and then draw a line around it. (Time, 3 minutes.)

In scoring, each item in each test is simply counted as right or wrong, and the total number of correct items on each test is considered the score on that test. The highest possible score is thus 46 points in addition, 30 in subtraction, and 24 on the other three tests. In making up a total score, or rating in total achievement in the fundamental subjects of the second grade, the scores on the four tests are simply added (the rights in addition and subtraction being added together and divided by two to prevent overweighting of arithmetic).[3]

[3] This rather crude method of obtaining a total on the scale was adopted only after considerable experimentation with more elaborate forms of handling. It has the prime merit of simplicity, and small likelihood of error in the doing. Transmutation of standing on each test into units of amount on a normal distribution curve was first tried. Values above the 95-percentile on each test were assigned a value of 7; values between the 95-and 85-percentiles a value of 6, and so on; these values were then added to give the total score. But the values thus obtained were found to correlate very highly (0.96) with values obtained by the crude method described above; the transmuting of values seemed, under the circumstances, not worth the labor.

Certain features of the form of the test remain briefly to be pointed out. The folder is of a convenient size—9 by 6 inches. The first page shows only the blank lines for the spelling test; and there is thus no opportunity for study of any test as the blank lies on the pupil's desk before the examination begins. Not over ten minutes are required for giving the spelling. Three minutes are allowed for each of the two reading tests, one and a half minutes for addition, and two minutes for subtraction. In fact, not over 25 minutes are needed for the entire examination—surely not an excessively long period even for the second grade. The examination is not wearying for the children to take, writing being required on the first test only. And scoring is easy and objective; for with a little practice an entire examination can be scored in one minute. In short, the effort throughout has been to make the scale as convenient and practical for survey work as possible.

III. The Development of the Scale

Work upon the scale began with a careful investigation of courses of study usual in the second grade. It was the aim of the writer to include in the examination only those subjects which are important in conditioning a child's school progress (the "promotion" subjects). Measurement of work in drawing or music was, therefore, not attempted. Writing was also omitted, because, though important, it was not a subject, according to the universal testimony of second-grade teachers, which could be considered critical—not a subject which actually did have any influence upon school progress. Reading was, according to the teachers, the most important subject in the second grade. This fact was recognized by devoting two tests to the subject. The other two fundamental subjects were spelling and arithmetic.[4]

For the spelling test, words were taken from the Ayres Spelling Scale, a list being made up after much trial which gave a fair measure covering the range of second-grade ability in the subject. The arithmetic tests are essentially similar to tests A and B of the Cleveland Survey series.[5] The arithmetic tests are distinguished

[4] It should be mentioned, in this connection, that the method of combining scores on the separate tests above referred to weights the subjects roughly in proportion to the teacher's estimate of their importance, the weighting running roughly 2, 1.5, 1 for reading, arithmetic, and spelling in order.

[5] Judd, Charles H., *Measuring the work of the public schools.*

from the other three tests (spelling, word recognition, and understanding of sentences) in that they are speed tests. The timing is carefully adapted to second-grade performance.

The two reading tests are, as far as the writer knows, original. The word-recognition test is the result of a long series of efforts at the development of a visual vocabulary test for use with young children. The problem was essentially a matter of directions. The test must be such that the directions could be so simple, natural, even colloquial, that a child would have no difficulty whatever in expressing, through the test, the extent of his reading vocabulary. The form described above was finally hit upon as filling, to a fair degree, these requirements. There were only ten zero scores out of 320 cases (0.3 percent); that is, failures to understand directions were very few. In fact, this game of finding the real word is not an uncommon practice in phonetics, and is readily entered into by the children.[6] The children's response is simple and easy; they merely draw a line around the real word. Further, scoring is the extreme of simplicity. Any indication of the correct word—a line around it, or a mark through it or any other reaction which points unequivocally to the correct response—is considered satisfactory. There is only one rule—that not more than one word should be marked in each line. The test of understanding of sentences, it may be said shortly, has similar advantages in giving, taking, and scoring; and it also has an extended history.

The preliminary experimentation was done in the schools of Bloomington, Indiana. A total of fifty-five words from the Ayres' list and fifty items each for the two reading tests were given to the Bloomington second-grade children. The difficulty of each item was expressed as distance on a P. E. scale; and items

[6] Indeed, another form of the test, with easier words, has been used successfully by the writer in the IB grade. It should also be mentioned that a distinctly useful test for the third grade might be formed by changing this test so that there would be four real words and one nonsense syllable in each line; the children would then be told to "cross out the letters that don't make a word." Such a form was experimented with but proved too hard for the second grade.

It should be mentioned that the words used in this Word Recognition Test were taken directly from the Haggerty Visual Vocabulary Test. (Haggerty, M. E. Scales of Reading Vocabulary for Primary Children. El. Sch. Jour. 17: 2, Oct. 1917.) The individual examination—to be referred to shortly—was therefore an examination with Haggerty's test.

giving fairly equal steps on the scale were incorporated into the final form. The arithmetic test differs from the preliminary form only in being somewhat shorter; the first form included all forty-five number combinations, but in the final form the list was cut down to forty addition and thirty subtraction combinations, the elimination being by chance. Shortly afterwards (in March, 1920) the whole scale was used in surveys of the school systems of Edinburg, Wabash, and Connersville, Indiana; and norms were established on the basis of data thus obtained.[7]

IV. Norms and Validation

Results from the three cities (a total of 320 cases) gave norms as follows:[8]

	Percentiles				
	10	25	Median	75	90
Spelling	3.5	7.0	12.2	16.6	19.1
Word Recognition	4.2	8.8	12.6	17.0	20.9
Understanding of Sentences	4.5	8.2	11.2	14.5	18.2
Addition	7.7	14.0	20.2	26.0	32.4
Subtraction	0.0	4.6	15.0	20.5	27.7
Total	28.3	39.6	52.3	66.3	82.0

The above standards will indicate, with some exactness, the scores to be expected of a second-grade class.[9] A question, how-

[7] The writer wishes to express her obligations to the superintendents and teachers of Bloomington, Edinburg, Wabash, and Connersville, for their kindness and cooperation in the progress of this work.

[8] The reader who is relatively unfamiliar with percentile distributions may better understand the practical meaning of this table if he considers that all the 320 scores in (let us say) spelling were arranged in order beginning with the smallest, and that when so arranged 10 percent of them were less than 3.5, 25 percent of them less than 7.0, 50 percent of them less than 12.2 (12.2 being therefore the median), 75 percent of them less than 16.6, and 90 percent of them less than 19.1. Similar meanings attach to the percentiles indicated in the table for the other tests.—*Editor.*

[9] The median age for the total group was 8.09; the examining was done the last of March. The median age is, the writer feels, distinctly important. All three school systems have an appreciable retardation. Other systems may have more, or less—that is, in other schools the children may be older or younger than the children from whom these norms were obtained. And it is essential that the maturity of the children be taken account of in comparing their records with the standards.

ever, arises as to the significance which may be attached to such scores with regard to school attainment. It remains to be proved, for instance, that the two "reading" tests really measure reading. A word remains to be said, that is, with regard to the validity of the tests.

About the validity of the spelling and arithmetic tests there can be little question; they are simple samplings of the abilities in question. That is, ability to add is measured by having the children do addition examples, and ability to spell is determined by having them spell. With respect to spelling it may be argued that the use of a list of words in this fashion is not a satisfactory test of spelling ability in a practical sense. But the point can hardly be made as far as the second grade is concerned. Spelling functions in the second grade only as "list spelling"; and the test furnishes this kind of an exercise. But, anyhow, such a theoretical consideration is really beside the point. Sentence spelling tests and timed tests are simply impossible in the second grade, because of the difficulty such young children have in writing.[10]

It remains to be proved, however, whether the reading tests really measure a child's ability to read. It is evident that they involve reading, and, *a priori*, they would be expected to give some measure of ability to read. But definite evidence in the matter is needed. Such validation has been secured in two ways; (1) by relating the scores on each test to carefully obtained teacher's ratings as to the reading ability of the same children; and (2) by relating the scores in the word recognition test to the scores made by the same children in recognizing exactly the same words when shown to them individually. (The examiner gave each child a very careful individual examination to determine the

[10] With the spelling and arithmetic tests, the only question is as to the reliability of the sampling. It may be said, for those interested in such matters, that the coefficients of reliability for these two tests are 0.92 and 0.98 for addition and subtraction respectively, and 0.88 for spelling; the probable errors are 1.19 for spelling and 0.99 and 0.96 for addition and subtraction. The coefficients of reliability for the two reading tests, it may be added, are 0.97 and 0.90; the probable errors 1.01 and 1.57.

The coefficients were calculated by using alternate items of each test to make up two half tests, correlating the two halves and applying Brown's formula. The P. E. was determined by finding the average difference between the two half scales and correcting according to the formula P. E. of the test $= \sqrt{2} \times 0.5978$ P. E. of the half scale.

extent of his acquaintance with these words.) These relations are best expressed by the correlation coefficient; and it may be said shortly that the correlation of the word recognition test with estimated reading ability was 0.73—a surprisingly high correlation for such work, and a striking validation of the test. The correlation of the test of understanding of sentences with estimated reading ability was 0.51—not so high, but a good correlation for a three-minute test with estimated standing. Both correlations are based on the total second-grade population of one of the cities surveyed (154 cases) and may be considered reliable.[11] The correlation of the word recognition test with the individual examinations was 0.82—and this, too, is a very satisfactory result. Thus the validity of the two reading tests is, as such tests go, surprisingly high.

In fact, all the tests of the scale seem to yield sufficiently consistent and sufficiently valid measures to make it trustworthy as an instrument for measuring attainment in the fundamental subjects of the second grade. Certain uses to which such an examination may be put may now be briefly discussed.

V. Uses of Such Scales

As was suggested in the first section of the paper, such a scale as has just been described should be of most obvious value for use in school surveys. It should give a means for measuring total achievement in the special subjects of the grade, make possible diagnoses of comparative strength and weakness in the various subjects, and yield a larger measure of the total pedagogical situa-

[11] In obtaining the teachers' ratings, a systematized rating scale was employed similar to the officers' rating scale used in the army. Comment should be made upon the unusual reliability of ratings in reading from teachers in the first and second grades. In these grades the child does his reading aloud, and every day he gives audible proof of his ability or lack of ability to read. The teacher has thus a very definite performance on the part of each child upon which to base her ratings. It is probable that this rating, being based on oral reading, tends to stress vocabulary rather than ability to understand the import of the material read. The lower correlation of the "understanding of sentences" test with estimated reading ability might be thus explained. Each one of the two reading tests apparently makes its own independent contribution as indicating reading ability. The intercorrelation between the two tests is 0.60; the partial correlation of word recognition with estimated ability, understanding of sentences value being constant, is 0.62, and similar partial correlation for the understanding of sentences test is 0.13.

tion in a given class or school than would be possible by employing merely separate tests in the various subjects. However, as was also mentioned in the first section, certain further uses of such scales have also been in mind. Two of these may be briefly indicated.

In the first place, such an examination should be capable of yielding a total score or "achievement total," which could be used as an indication of a child's total status in the fundamental work of the grade in which he is placed. The concept is obvious enough, and need not be elaborated upon. It should only be pointed out that wherever school work is organized according to grades, and not departmentalized according to subject, the grade, and not the subject, is, in many important respects, the unit. Therefore an examination fitted to give a statement of total attainment within the grade would seem the natural type of measuring instrument to use in dealing with many problems which come up in grade-school teaching. At least, such a scale for such uses—e.g., in making section divisions—would seem well worth trying.

In the second place, an examination of this type, if used in conjunction with a group scale of intelligence, should give a very valuable means of measuring the efficiency of teaching. Without such supplementation it does *not*. A fair measure of the work of a teacher cannot be obtained by measuring merely the level of attainment to which she has brought her children. This can only be done by considering what the class has done in relation to what it was capable of doing—by setting over against each other achievement and ability. When such double measures *are* obtained, the information they yield with regard to a teacher's efficiency is sometimes striking. Two second-grade classes in a system recently surveyed by the writer will serve as examples. Both the examination described in the present paper and a group scale of intelligence[12] were given to these two classes. One school was found to have only 22 percent above the median for the entire city in attainment; 33 percent of the children were above the median in intelligence. Apparently this is a school in a less desirable district, and (as often happens), it is not the district to

[12] Pressey, L. W. "A group scale of intelligence for use in the first three grades," *Journal of Educational Psychology*, 10:297–308, September, 1919.

which the best teachers are assigned. The second school shows 74 per cent above the median in attainment, and 62 percent above the median in ability. The teacher in the second school, we may infer, is unusually able. We might, in fact, throw the figures into percents, and say that the first teacher was only 67 percent efficient (since the standing of her class in the school subjects was only 67 percent of its standing in innate ability). By the same method the second teacher would then be regarded as 120 percent efficient. The exact form in which the results are expressed is immaterial.[13] But the notion of thus evaluating teaching by measuring the attainment of a class of children with reference to their ability is, the writer feels, decidedly worth while.

Summary

The paper may be briefly summarized:

1. The study describes the attempt to combine in a single examination such tests as would yield a summary statement of standing in the fundamental subjects of the second grade.

2. The scale developed consists of four tests, designed to measure abilities in spelling, reading vocabulary, silent reading, and addition and subtraction. The tests are presented, one on each page of a four-page folder, the entire sheet being 9 by 12 inches. The examination takes 25 minutes to give; and the blanks can be scored at the rate of about 60 an hour.

3. The tests have been carefully developed and the items arranged on a P. E. scale. Tentative norms are presented, based on a total of 320 cases—all the second-grade children in three small cities.

4. Data are presented which show the test to be reasonably reliable and valid.

5. It is suggested that the scale should prove useful for the following three purposes: (a) it should be decidedly valuable for use in educational surveys; (b) it should aid in the educational readjustment of individual children by yielding a summary statement of their standing in the fundamental subjects of the grade; (c) when used in conjunction with a scale of intelligence it should give a very valuable index of teaching efficiency.

[13] It should be confessed at once that the forms used above are not statistically by any means all they should be; but they will serve to denote the general nature of the contrasts.

Editorials

A BAD SITUATION MADE INTOLERABLE

The teacher shortage is not new. It has always been with us, but it has been masked by the willingness of boards of education and trustees to employ relatively incompetent people. When the Commissioner of Education declared, as he did at the Citizens' Educational Conference last May, that 120,000 teaching positions would be vacant in September with only 30,000 prospective teachers coming from normal schools and colleges, he indicated an alarming condition. Half or more of the remaining 90,000 positions will doubtless be occupied if not filled. The doors of thousands of schoolhouses will be open, and within them certain gestures will be made more or less resembling those of teachers. Customary signs of activity will appear, and the passer-by may derive no small satisfaction in contemplating them. As a wayfaring man, though he may be no fool, he will, however, "err therein." For no more insidious fraud is beguiling the public into misplaced confidence than this mimicry of true teaching. A real educational lock-out is far better, for with closed doors and pupils turned away, no one can be deluded into supposing that the constitutional rights of children are being assured.

So it has been in other Septembers, and so it will be in Septembers yet to come. The discrepancy in other years may not be 90,000; but there has been in the past and there will continue to be in the future a shortage less critical perhaps than the present one but nevertheless too real to be anything but ominous.

It appears to be nobody's business to remedy this calamitous condition or even to know the facts in the case. A prominent official in a certain state—a man who ought to know the real conditions if anybody does—stated less than a year ago that the schools of his state "lose 3,000 teachers each year and therefore need 3,000 recruits annually." In the state in question the true number of recruits needed annually is certainly much greater. It

may be as many as 5,000. Indeed, only by supposing that the average length of service for each teacher in that state is over thirteen years, could it be true that only 3,000 recruits would be needed. If the average period of service is actually more than half of thirteen years it is exceptionally long. Moreover, we cannot equate "teachers lost" with "recruits needed" except on the assumption that school conditions are fixed—on the assumption, in other words, that no new positions are created. This would mean no extension of educational service to include new special classes, no additional supervision, and no increase in population. In many states these changing conditions will make "recruits needed" larger by several hundred than "teachers lost." If the official to whom we have alluded manifests the prevalent state of mind concerning the number of teachers needed, we have one reason for perennial shortage.

Because of ignorance of conditions on the part of those who ought to be informed, the teacher-training establishment has in most states no visible relation to the demand for teachers. Even if all the graduates of state-supported normal schools and schools of education became teachers in the state at whose expense they were trained, they would not suffice to fill half the required positions. It is notorious, however, that they do not all serve the state as teachers. In a certain state—one far above the average educationally—it was found that not more than eight out of ten did so. Of the other two, one never became a teacher and the other went outside of the state to teach. Even the eight who served the state did so, in most instances, for but a short time. One was lost at the end of the first year, another at the end of the second year, and a third at the end of the third year. Thus in the state to which we refer not more than five out of ten of the graduates of state supported teacher-training institutions served the state longer than three years. The average length of service of all graduates who become teachers was less than five years. And all this was before the recent crisis.

Thus the demand is enormous, and supplying it is like pouring water through a sieve. Consider a state whose teaching force is 20,000. Such a state would in normal times need about 4,000 new teachers annually. It might, if its salary conditions were normal, expect to import about 400 teachers a year, and if it were

reasonably well provided with privately endowed colleges it might obtain 300 more from that source. It would have to provide from publicly supported teacher-training institutions for about 3,300 positions. This would require quite 4,000 graduates—less, of course, a certain unknown number of certificated teachers of less training. Ten or twelve states have 20,000 or more teachers, and most of them have use in normal times for upwards of 4,000 new teachers a year. In the past none of them has been providing that number of recruits from normal schools and colleges, and at present the methods hitherto used have become pitifully inadequate. The shortage was very real before it assumed its present proportions. The situation has been bad for a long time; it has now been rendered intolerable.

Nevertheless the present crisis brings its advantages. A new sense of values is being developed, and a certain detachment from precedent has been won. It has become clear that, on the basis of standards already accepted in business, teacher-training as a state enterprise is astonishingly ineffective and haphazard. Kindergartners trained as such at state expense are taking graded-school and even high-school positions—and that too when there are more kindergarten vacancies than there are kindergartners to fill them. In a certain state a new course for training domestic-science teachers was established, although the number already being trained exceeded the demand. Students trained for the elementary schools are becoming high-school teachers, despite the fact that they are needed in the service for which they were trained.

It should be somebody's business to know these conditions and to take action with reference to them. In each state a bureau should be established which should analyze the types of teaching service, determine the demand for teachers in each of them, and provide training facilities to meet the demand. Such a bureau should have at its disposal a large number of scholarships which might be awarded now to one type of students and now to another according to the need of the schools for teachers of one kind or another. It should obtain and file records concerning each teacher in service in the public schools and concerning all admissions, withdrawals, and transfers affecting the service. It should have supervision of the placement of graduates of teacher-training

institutions and should see to it that the product of these institutions is delivered to the public schools. In short, it ought to be somebody's business to study not sporadically but continuously the large problems of training, placing, and holding a sufficient corps of competent teachers. It is equally clear that such an authority should have power to put its findings into effect.

B. R. B.

WHAT TESTS ARE FOR

Everywhere school people who are not hopelessly numbered with the professionally moribund are becoming interested in the uses of intelligence tests. Within the last year something like ten serviceable intelligence scales for group testing have been published, covering the ranges of ability from kindergarten to university. We may expect that within the next twelvemonth several million school children will be tested, and that the number will rapidly increase from year to year. We may be perfectly certain that wide-awake principals and superintendents will discover them to be one of the most valuable tools ever placed in the hands of the school administrator. They do not yet know, no one knows, all the uses which will ultimately be made of this tool; but with the aid of our research directors in the schools we may expect them, through well-planned trial and careful checking-up of results, to find out.

But just as it is impossible to classify all children as sick or well, tall or short, normal or feebleminded, so it is impossible to classify all school principals as either scientifically progressive or hopelessly conservative, that is, as alive or dead. Many are only in a semi-comatose condition, suffering, it may be, from the results of intellectual malnutrition or the deadening effects of administrative routine. One whose training has been inadequate, or who spends his life in an atmosphere saturated with the poisonous vapors of tradition and routine, needs a good deal of professional vitality to keep pace with educational developments in these days of standardized educational tests, intelligence scales, vocational guidance, supervised study, project teaching, *et cetera*. It is not strange if some suffer noticeably from mental confusion or drop occasionally into a light slumber. Such are not necessarily hopeless. They are easily stimulated into vigorous life by the din of a

new educational battle cry. Tests and measurements are having this effect.

It must be admitted, however, that some are aroused only sufficiently to enable them to go through the motions of participating in the new movement. If it is a question of intelligence tests, these are given, but the results are not used to effect any improvements in grading, classification, instruction, or educational guidance. Nothing is gained beyond the satisfaction of being able to say that children in such and such number have been tested and that such and such interesting facts have been discovered. It is true that even this may have a certain value in alleviating the boredom of tired school people, but it seems necessary to insist upon the fact that tests and scales are designed to serve, not as playthings, but as serious tools.

It is not the purpose of this editorial to try to teach the use of such tools. A manual, of a kind not yet written, would be needed to accomplish that. Meanwhile, any resourceful principal or superintendent who gives serious thought to his problems will be able, in the light of his test results, to devise helpful plans for adjusting the curriculum, the grading system, and the methods of instruction to the individual differences of pupils. Don't play with tests; use them.

L. M. T.

Reviews and Abstracts

KLAPPER, PAUL (Editor), *College teaching*. Yonkers-on-Hudson: World Book Company, 1920. 583 pp.

It is rather generally admitted that good teaching—I mean teaching as an art—is not as often practiced in higher educational institutions as it is in secondary and elementary schools. There are many reasons why this is so. Some of them are good reasons; others are not. The greater maturity of college students and their relative independence of the teacher, the more commanding knowledge of subject-matter which the college teacher must possess, the lack of any means for training college teachers as high-school teachers are trained—these are among the good reasons, their goodness being inversely in the order of mention. The appointment of college teachers with little or no reference to teaching ability, their promotion in rank with the same disregard of skill in instruction, and their recognition among holders of similar positions throughout the country not because of their eminence as teachers but because of their success as investigators—these are bad reasons, their badness being due not to their lack of cogency as reasons but to their undesirability as facts.

This book is apparently offered as a first step toward the improvement of this condition. The editor says: "This volume is conceived in the hope of stimulating an interest in the quality of college teaching and initiating a scientific study of college pedagogy." Since it was thought that "no college teacher. . . . can speak with authority on the teaching of all the subjects in the college curriculum," the plan contemplated the treatment of each of fifteen important subjects by a different author. The contributors were previously furnished with an outline of which the following were the major topics: (1) aim of the subject; (2) its place in the college curriculum; (3) its organization (sequence of courses, electives, etc.); (4) discussion of methods of teaching the subject; (5) moot questions in the teaching of the subject; (6) tests—e.g., of the worth of the subject to the student, of the accomplishment of the aims, and of how much the student carried away; (7) bibliography of the pedagogy of the subject. A good outline, I think; and one which, if it had been followed, would have gone far to realize the editor's hope of an aroused interest in college pedagogy.

But it was not followed. Only a few of the contributors made any real attempt to follow it, and a majority of those who did missed the point. It is clear from the editorial preface and from the "Introductory Studies"—especially those by President Mezes and by Professor Klapper—that the central idea in projecting this volume was "methods of teaching." Yet it is to administrative questions that most of the writers of these chapters devote their space—to the value of their subjects, the place of each subject in the curriculum, and the number of prescribed and elective courses. They have little to offer on method. They usually distinguish the lecture method, the recitation method, and—where it applies—the laboratory method. Their observations, even on these mechanically obvious types of teaching, tend to be platitudinous.

Of the opportunity afforded by their subjects for inductive and deductive work, drill and review, study lessons, lessons devoted to appreciation, to say nothing of project lessons and socialized recitations—of these opportunities they have almost nothing to say. The recognized steps in teaching (preparation, presentation, and others less universal), the technic of questioning, the value of planning—in short, most of the concepts and processes that enter into teaching *as such* are not recognized. Yet most of them have as obvious a bearing on college teaching as on high-school teaching.

Although these writers are rather more likely than the general run of college teachers to be sympathetic to the technic of teaching, some have nothing more to offer than a general sympathy, because they have not themselves been trained in education. Others evidently hold to the notion that "the only equipment for successful teaching is a thorough knowledge of the subject." Professor Lemon in the chapter on *physics* declares: ". . . . the teacher must first of all know and understand his subject. Right here lies probably nine-tenths of the fault with our pedagogy. No amount of study of method will yield such returns as the study of the subject itself." Such ideas do not take account of the fact that the scholarship of hundreds of physicists, chemists, psychologists, and linguists does not save them from being notoriously poor teachers.

This attitude, however, is not always taken. Doctor Neumann's treatment of method, for example, in his chapter on *ethics* is entirely competent. Professor Fetter gives a good account of several general methods of teaching *economics*. Professor Galloway has some thoughtful paragraphs on laboratory teaching in his chapter on *biology*. Some practical notes on method in teaching *engineering subjects* are given by Professor Baker. And Professor Winchester's ideas on teaching appreciation in *English literature* are decidedly good.

Besides, there are several chapters which have decided interest despite the fact that they miss the point. One's attention would be at once arrested by Talcott William's essay on *journalism*. Professor Miller's chapter on *mathematics* has dignity and breadth of view. The best written chapter is the one by Doctor Canby on *English composition*. It is the best thing of the kind with which I am acquainted, except Barrett Wendell's "English Composition." The classification of composition teachers into four parties—the Know-Nothings, the Do-Nothings, the Formalists, and the Optimists—is both clever and convincing.

If the purpose of this book was to show not how the methodology of education is actually being applied in college teaching but how it might and should be applied, the purpose would have been better accomplished by a single educational expert with the assistance of specialists. The admirable chapter written by Professor Klapper on the "General Principles of College Teaching" shows that this could be done. But if, on the other hand, the purpose was to give an accurate impression of the way college teachers are thinking about teaching, this device of multiple authorship was precisely the way to serve the purpose. The real value of the book lies, in my judgment, in its exposition of contemporary thought in an important field of teaching. Such an exposition is decidedly important, and no other book, and not many combinations of books, will give it as well as this one.

B. R. B.

TRABUE, M. R. and STOCKBRIDGE, FRANK PARKER. *Measure your mind.* Garden City, New York: Doubleday, Page and Company, 1920. 349 pp.

The use of psychological tests in the army and in an increasing number of industrial establishments has served to bring such tests to the attention of a large and diversified audience. It is not, therefore, surprising to receive a book upon psychological tests which is addressed "to employers and those in charge of the selection, grading, and promotion of workers of every class in factories, offices, and stores; to teachers of all grades, from kindergarten to university; to parents who are interested in ascertaining, and watching the growth of their children's mental development and advancement and desirous of learning something of their own mental capacities and limitations as a guide to the intelligent choice of vocations or professions," and devoted to a general description of mental tests and their use.

In the second part of this volume thirty mental tests (called Mentimeter Tests by the authors) are reproduced and the procedure to be followed in administering them is described in detail. In this list one finds a few tests that are new and a number of familiar ones that have been modified slightly for this particular purpose. They are grouped as follows:

I. Tests for Infants.
II. Group Tests for Non-English-Speaking Persons.
III. Individual Tests for Non-English-Speaking Persons.
IV. Group Tests for Persons that Read English.

The authors take the position that a much greater service is rendered by giving a large list of tests and accompany this list by advice for the selection of a group of tests for particular purposes. Such advice is given in one chapter of the book. Preliminary or tentative standards are given for each of these tests which are equivalent to (1) inferior ability; (2) low average ability; (3) average ability; (4) high average ability; (5) superior ability.

The combination of authorship and the general character of the book indicate that it was the intention of the authors to present in a practical and semi-popular manner the technical topic of psychological tests. In this endeavor they have been partially successful. The language of the book in some cases, although simple as to vocabulary, is involved as to sentence structure. The authors also failed to make use of the device of topical headings; the book is not accompanied by an index.

W. S. M.

WOODROW, HERBERT. *Brightness and dullness in children.* Philadelphia: J. B. Lippincott Company, 1919. 322 pp.

A letter of Francis Galton written the day before his fifth birthday is dramatically contrasted with the case history of Abbie, an inmate of an institution for the feeble-minded. Francis "can read any English book" and "can say all the Latin substantives and adjectives and active verbs, besides 52 lines of Latin poetry." He can "cast up any sum in addition" and multiply by the numbers from 2 to 11. He knows his "pence table," can "read French a little," and can tell time. On the other hand, Abbie after ten years of courageous instruction has reached her limit. She is twenty-two years old. She can braid corn husks a little, make a bed, and iron an apron. She can

read a few sentences, spell a few words, and write about twenty-five words from memory. She knows the days of the week but not the months of the year and does not know how many fingers she has on both hands. She is about an average feeble-minded person.

Between Francis and Abbie are the children who attend our public schools. From brightness approaching that of Francis with all its wonderful opportunities, to dullness like that of Abbie in which twilight broods over intelligence, Tom and Dick and Harry, Susie and Mary and Jane find their peculiar places. Children of the same age and schooling vary from one to five years in mental capacity, and those who differ from "normal" are more numerous than we suppose.

It is the children of low intelligence upon whom we have centered our attention. To the fact that they are easily identified and are incumbrances in the grades has been added the philanthropic idea of improving their unfortunate condition. Bright children are not so easily identified, are usually a joy to the teacher, and are too self-reliant to appeal to merely sentimental notions about their amelioration. Yet the hope of improving the feeble-minded can be realized only in a very limited way, while, every effort made to give bright children the fullness of life of which they are capable yields gratifying results.

That teachers need a knowledge of the modern psychology of intelligence is the contention of the author, and this book is intended to impart the knowledge. It seems well calculated to do so. It presents, for example, in a simple but effective way the rationale of the Binet-Simon Scale and the evidence of its accuracy as a measuring instrument. On the latter point the author says: "There is scarcely an important criterion of accuracy with which the tests do not comply." Accepting as the author does the construction of the Binet-Simon Scale as the most valuable contribution of psychology to education, he has little to say about the possibility of group testing. The Binet Scale is indeed a "wonderful achievement" but its general use is precluded because of the expense of administering it.

One of the strong features of the book is the care with which its terminology is defined and used. The distinction between mental age and brightness, the meaning of the terms idiot, imbecile, and moron, the difference between chronological, pedagogical, anatomical, and mental age—these are carefully formulated. These formulations have, of course, become much more practicable since the development of a means of measuring intelligence. Thus idiots are defined as "those so deeply defective that their mental development does not exceed that of a normal child of about two years," imbeciles as those whose mentality does not exceed that of a normal child of seven years, and morons as those who grade between seven and twelve years.

The chapters on the relation of physical condition to intelligence are well balanced with reference to the other parts of the book. Their treatment is non-technical, as indeed it should be, and some of the deductions are of great significance. The fact that no child ever becomes feeble-minded after the age of four or five years is brought into significant relation with the fact that the brain reaches almost its full weight at that age. Again, based on the

physical correlate of mentality, the author concludes that the really critical years of a child's life are the years before he enters school. The age of puberty —often regarded as particularly critical—"is incomparably less critical than the first few years of life."

But what is intelligence, or rather, what is its index? For Ebbinghaus it was the combining activities of the mind; for Wundt the ability to give attention or to concentrate; for Binet a sound judgment; and for Stern adjustment of thinking to new requirements. While it is plainly impossible to find any one mental process that can be identified with intelligence, many abilities have been measured and their reliability as indices has been studied. The author takes up at some length these various capacities. They are examined under the heading of simple sensory capacities, more complex capacities such as association, memory, and attention, and complex mental processes.

The treatment of these traits and capacities consists largely in a summary of the ways that have been devised for measuring them, and the extent to which each of them when so measured correlates with general intelligence. In these chapters the author departs from the helpful treatment of his subject which he exemplifies at the beginning and again toward the end of his book.

After taking up certain questions of heredity and pronouncing the dictum "that heredity rather than environment is the preponderating factor in the causation of individual difference in mental capacity" and after supporting this contention by some interesting evidence as to the heredity of feeble-mindedness and human eminence, the author devotes two exceedingly good chapters to practical implications for school procedure. "Granted," he says, "that original abilities are inherited, they are not, however, inherited in their final form." "The business of the schools is to see that the child has the chance to show what is in him." Practical suggestions follow as to what may be taught and what may not profitably be attempted to be taught to children having specified amounts of mental defects. Similar suggestions are offered with reference to gifted children. Here he makes a strong plea for the importance of this type of work. He says: "If we continue to use a large portion of our energies in training the children of the lower levels of intelligence without simultaneously devoting special attention to the exceptionally bright we shall produce an unbalanced system which must in time tend to pull downward the general level of education."

It is our opinion that this book is entitled to receive and will receive a wide and appreciative reading.

B. R. B.

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Walter S. Monroe, University of Illinois, Urbana, Illinois.

Wisconsin Association for Educational Research

The organization of the Wisconsin Association for Educational Research indicates the rapidly growing interest of the school men of that state in the study of quantitative as well as qualitative problems of education. The personnel of the association includes representatives from practically every field of educational work in the state. The charter membership includes teachers, supervisors, superintendents, and members of normal school and university faculties who are studying educational problems.

The purpose is not only to encourage the scientific study of educational questions but to keep the members closely in touch with what each is actually doing in the way of educational investigation and experimentation. By means of a bulletin issued at regular intervals the members can announce intensive studies and ask for co-operation or for the results of similar studies completed or under way by other members in the same or similar subjects. They can learn definitely and economically to whom to go for specific information and can avoid a repetition of unnecessary effort.

The state department of education in Wisconsin is well represented in this group. Dr. W. W. Theisen and Mrs. Cecile White Flemming, supervisors of educational research in the state department, have been influential in the formation of the organization. The officers for the year 1919–1920 are Supt. C. J. Anderson of Stoughton, president; Mrs. Cecile White Flemming, state supervisor of educational measurements, secretary-treasurer.

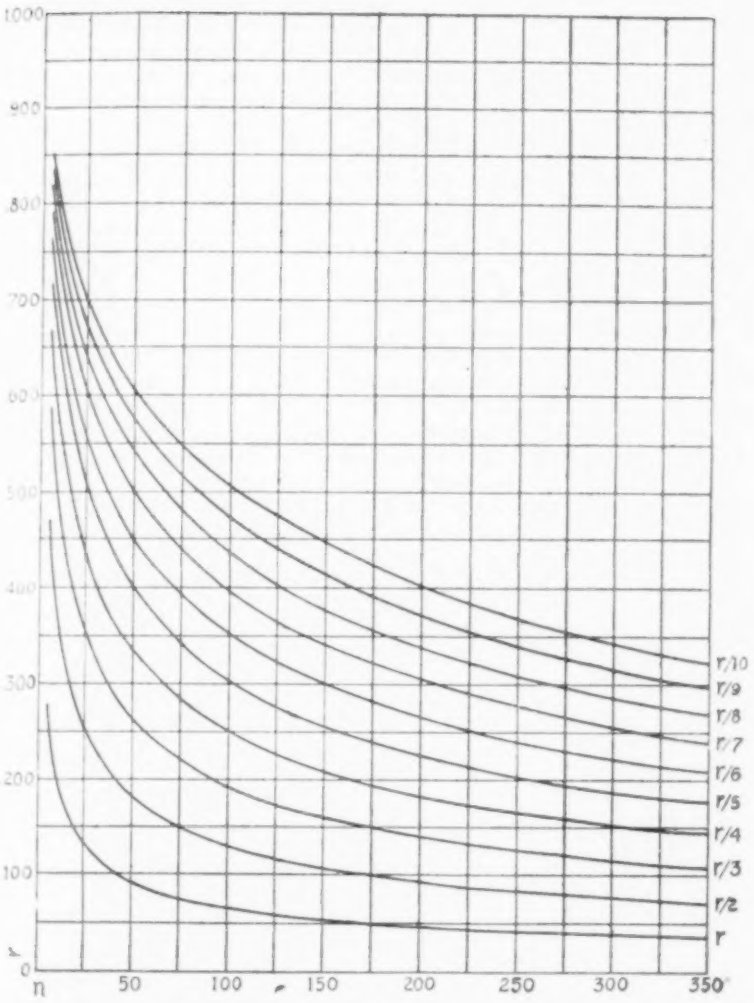
A Short Method of Determining Probable Error

It is the purpose of this communication to present an easy method of determining the relation of the probable error to the Pearson coefficient of correlation.

The formula P. E. $= .67449\frac{1-r_2}{\sqrt{n}}$ has been solved previously for many values of r and n, and the results have been arranged conveniently in tables, such as those found in Rugg's *Statistical Methods Applied to Education*, page 404.

Referring to such tables the value of P. E. is determined and its relation to r is noted. Since its relative value with respect to r is of greater importance than its absolute value, an easy method of determining this relationship was considered desirable.

The method was quite simple. It consisted in substituting, first, P. E. $= r$ in the formula, and then solving for r through the various values of n from 5 to 350. The same process was repeated with P. E. $= r/2, r/3, \ldots r/10$. By plotting values of n on the abscissa, and values of r on the ordinates of the accompanying diagram, the several curves of P. E. were plotted.



By referring to this diagram it can be determined by inspection that if $r=0.250$ and $n=100$, P. E. $=r/4$, or P. E. (if one wishes the absolute value) equals 0.63. Again if $r=0.300$ and $n=50$, P. E. lies between $r/3$ and $r/4$. Since these ratios determine the reliability of r, such an easy means of determining their values has proven very useful to the writer.

H. J. Baker,
Assistant Director, Psychological Clinic,
Detroit Public Schools.

Some Correlations with the Otis Scale

Through the cooperation of five teachers in the eighth grade of the Wellington city schools the data for this study were secured at the end of the school year 1918–1919. Seventy-six pupils took the Otis Group Intelligence Scale. In addition, independent ratings were made by the five teachers with instructions to consider only general intelligence and ability, without regard to classroom achievement. Letter ratings were used under the following restrictions. The highest rating should be given to eight pupils or approximately 10 percent; the next highest rating should be given to 15 or approximately 20 percent; the median rating to 30 or approximately 40 percent; the next to the lowest rating to 15 or approximately 20 percent; and the lowest rating to 8 or approximately 10 per cent. The records of the school also furnished the year's grades of the 76 pupils in the subjects taught by the five teachers who gave the intelligence ratings.

The Pearson coefficient of correlation between the teachers' ratings and the grades issued by the teachers was 0.882, between the teachers' ratings and the Otis scores 0.598, between the school grades and the Otis scores 0.582.

The more significant results are found in the correlation of the teachers' ratings with each of the separate tests in the Otis test and in the correlation of the total scores in the Otis test with the scores made in the separate tests.

Otis Sub-Test	Correlations with	
	Teachers' Ratings	Otis Total Score
1	.587	.598
2	.259	.585
3	.029	.432
4	.415	.811
5	.425	.510
6	.511	.642
7	.343	.639
8	.316	.531
9	.583	.390
10	.368	.510

It is, of course, obvious that the number of cases here used as a basis for study is insufficient to warrant any final conclusions. It would seem, however, that the various tests of the Otis Scale are of different value. This suggests the possibility of deriving a shorter test for school uses.

A. K. Loomis.

Principal, Summer County High School, Wellington, Kansas.

The Bureau of Educational Research at Aberdeen South Dakota

The Bureau of Educational Research of the Northern Normal and Industrial School was established in 1917 for the purpose of aiding the schools of the state in studying scientifically certain types of school problems. Most of the work of the Bureau has been that of giving and scoring well-recognized, standardized tests. At first these tests were given in the schools and sent to the Bureau of Educational Research for scoring. This placed a great burden upon the Bureau and there was much delay in getting results back to the teachers. Furthermore the director of the Bureau, M. R. Staker, has always felt that standardized tests would accomplish the best results only when the teachers out in the schools did their own scoring and made their own interpretations. If the tests are valuable, they should be handled by the teacher and not by some student of education who is gathering material for a thesis. Perhaps it is true that any new movement must start with the investigator gathering data for theorizing, but the time has now arrived when educational tests should be put into the hands of the teacher in order to obtain the very best results from their use.

This Bureau is attempting to render service in the following ways:

1. The Bureau is offering a thorough course in educational measurements to regular normal school students in attendance. In this course the student is given an understanding of all the statistical methods needed in the administration of educational tests. Standardized tests in all the subjects are studied; students are given practice in giving these tests to children, in scoring test papers, and in suggesting possible remedies for discovered weaknesses.

2. The Bureau is keeping on hand a supply of well-recognized tests which are sold to the schools at cost. Instructions and any help teachers need in using these tests may be had by simply asking.

3. The Bureau realized that superintendents often wish to do standardized testing but find that their teachers are not sufficiently trained. The Bureau therefore has offered to give teachers a thorough course in educational measurements. A superintendent may act as conductor for a study group and the Bureau will send lessons and other helps to the group. The work has been so carefully planned that even if the superintendent himself is unfamiliar with standardized testing, he can study this with his teachers. Teachers completing this course are required to give tests to their own pupils, work up the results in all possible ways as suggested by the Bureau, and send the conclusions to the Bureau of Educational Research for correction. At present the Bureau is conducting this work in twelve study centers. It is possible for a group to finish this course in about ten weeks. After the group has finished, the superintendent has an enthusiastic group of teachers ready to help him carry on scientific testing. The Bureau wishes that county superintendents would take this work up with their rural teachers. So far only a few superintendents have undertaken the work, but surely more will do so next year because it is an established fact that the Bureau can carry on this work through the study center. It is the plan of the Bureau to continue this method of reaching out to the schools until every child in this vicinity of the state shall feel the effects of its work.

The director of the Bureau has compiled a brief treatise on the statistics of educational measurements which attempts to make clear the hidden mysteries of this

subject. In other words this little text talks in a simple way about a simple thing to the teachers. The work has become so simple that very seldom is the Bureau required to make any further explanations. Any educator, wishing a copy of the little text on the statistics of educational measurements by M. R. Staker, may secure one by sending $1.00 to the Bureau of Educational Research, Northern Normal and Industrial School, Aberdeen, South Dakota.

Educational Research in Cities of Over 10,000

In order to find out the status of educational research in the city school systems, a letter was sent in January, 1920 to the superintendents in cities having a population of 10,000 or more. The letter was not sent, however, to cities where a member of the National Association of Directors of Educational Research was in charge of this type of work. Letters were also sent to a few cities of less than 10,000 population where there was special reason for thinking that research work had been provided for.

In the circular letter the following information was requested: (1) Has provision been made for the systematic use of standardized tests or other scientific methods in your school system? (2) Name the official title of the person in charge of the work if such provision has been made.

Twenty-two cities in the United States have persons in charge of research work in the public schools who are members of the National Association of Directors of Educational Research. In addition to these an affirmative answer to the first question was received from 159 cities out of a total of 330 replies. In 108 of the 159 cities the superintendent was named as the person in charge of the work. In the remaining 51 cities some person other than the superintendent was named. Frequently the official designation of this person indicated that he had other responsibilities. It is, however, very significant that in so many cities this research work in the public schools has been recognized to the extent of making definite provision by the designation of some person other than the superintendent to have charge of the work.

A summary of the reports by states is given in the table below. In addition to the 159 cities which reported provision for the systematic use, 72 others stated that some form of research work was carried on irregularly.

State	Number Sent	Replies Received	Systematic Use	Irregular Use of
Alabama	8	2	0	1
Arizona	4	2	0	0
Arkansas	9	5	3	2
California	23	10	4	5
Colorado	15	9	6	2
Connecticut	30	9	5	3
Washington, D. C.	1	1	0	0
Florida	7	1	1	0
Georgia	19	4	0	1
Idaho	7	3	2	0
Illinois	38	20	11	2

State	Number Sent	Replies Received	Systematic Use	Irregular Use of
Indiana	34	11	5	3
Iowa	20	5	4	0
Kansas	21	10	9	0
Kentucky	9	5	2	1
Louisiana	7	3	2	1
Maine	10	1	1	0
Maryland	3	1	0	0
Massachusetts	62	31	6	10
Michigan	33	15	9	3
Minnesota	11	7	4	1
Mississippi	7	5	2	0
Missouri	19	10	5	2
Montana	7	3	2	1
Nebraska	3	2	1	0
New Hampshire	8	4	1	2
New Jersey	36	18	10	5
New Mexico	3	2	1	0
New York	51	21	7	4
North Carolina	13	2	1	0
Ohio	46	17	10	4
Oklahoma	9	7	3	2
Oregon	8	4	2	2
Pennsylvania	67	30	19	3
Rhode Island	9	4	0	1
South Carolina	6	2	0	1
South Dakota	5	1	0	1
Texas	25	6	0	4
Utah	4	2	0	0
Vermont	4	2	1	0
Virginia	10	6	3	0
Washington	12	5	5	0
West Virginia	9	4	1	2
Wisconsin	21	14	7	3
Wyoming	5	4	4	0
Total	758	330	159	72

National Association of Directors of Educational Research

(E. J. Ashbaugh, *Secretary and Editor*)

The members of the National Association of Directors of Educational Research are both busy and modest. The secretary has been impressed with this fact again and again. At different times a general invitation has been extended for news items and contributions which would tell of the good work being done. Very few answers have been received. But—a hurry-up call in midsummer for specific information on the following points: summary of testing activities during May and June; plans for the coming year; teaching activities of the summer; and possibility of attending the next annual meeting next February brought 34 replies in ten days. And the testimony of testing work done, to say nothing of the other lines of research which was undoubtedly done but not mentioned, was the most positive evidence available of the first statement in this paragraph.

Of the thirty-four replies received before these items were written, seventeen members stated that they expect to attend the next annual meeting at Atlantic City; four more stated that they hoped to do so, and three that they were undecided. Four stated that it would be impossible and six did not answer that particular question. Since a number of those from whom no replies were received reside in the east, the prospects at this early date look good for a large attendance at that time.

Eighteen of the thirty-four taught one or more courses allied to the work of the Association in various universities and normal schools this summer. The following taught in their own institutions:

De Voss, Kansas State Normal School, Emporia, Kansas.
Harlan, State Normal School, Lewiston, Idaho.
Staker, Northern Normal and Industrial School, Aberdeen, S. Dakota.
Buckingham and Monroe, University of Illinois.
Greene, Horn, and Ashbaugh, University of Iowa.
Fordyce, University of Nebraska.
Gray, University of Chicago.
Kelley, T. L. and McCall, Teachers College, Columbia.

In addition, the following are also teaching:

Alexander, University of Chicago.
Anderson, University of Michigan.
Courtis, University of Chicago.
Lane, University of Southern California.
Smith, Kearney State Normal School.
Theisen, University of Wyoming.

There are doubtless many others who are spreading the good news and inspiration of educational research work, especially in the fields of measurement and accounting, in other universities and normal schools. And these who are so doing are not all

members of the Association by any Means. The teaching profession is undoubtedly rapidly becoming aware of some of the advantages attending the use of these instruments.

An examination of the educational tests given during the spring of 1920 reveals a long list which has been used in one or more places. Most noticeable in terms of the number of tests used and the number of places using them (on the basis of these thirty-four replies) are: Courtis Series B; Monroe Silent Reading and Spelling, the last of which seems to have been in terms of various tests made from the Ayres Scale, Ashbaugh-Iowa Scale and Jones Demons. Among those mentioned less frequently were: Arithmetic: Woody; Woody-McCall; Boston Fractions; Cleveland Survey; Monroe's Diagnostic; Stone's Reasoning; Buckingham's Reasoning; and the Monroe Reasoning. Miscellaneous tests mentioned one or more times were: Charter's Language; Greene's Organization; Thorndike's Alpha 2; Van Wagenen's History; Brown's Latin; Henmon's Latin; Trabue's Language; Courtis' Geography and general writing tests.

Among the intelligence tests mentioned were those by: Haggerty; Otis, Pressey; Johnson; Van Wagenen; and the Army and the National Intelligence Tests. The following plans for next year in this line will give some indication of the tendency toward the use of Intelligence testing.

Pittsburgh expects to use from sixth grade through high school, Boston in grades VI-VIII in two schools; Detroit in first grades; St. Paul in seventh and eighth and to selected children; Santa Ana, Cal. in grades I to VI with the Otis and National Intelligence Tests in grades VII-XII; Beaumont, Texas in grades I and IX for classification purposes. Omaha and Leavenworth, Kansas also have large plans for the use of such tests but are less specific concerning them.

Other types of activities of the members which may be of interest to the membership of the Association are reported as follows:

Richardson, who reported on "Campaign Method in Elementary Education" at the Cleveland meeting and whose paper has appeared in this Journal, announces his plan for a similar campaign this coming year in spelling. Perhaps the way to teach spelling most economically is by an intensive campaign for a few weeks each year only. Who knows that it is not?

Anderson at Detroit reports plans for testing and developing tests in typewriting, chemistry, algebra, and foreign languages. We shall expect something of superior worth to come forth within the next two or three years as a result of these plans.

Talbert reports that he hopes to develop practical ability (both general and specific) tests for use among adults in industry and business.

Alexander is continuing his work on application of salesmanship and advertising discoveries to school publicity. He is also making a study of the secondary school teachers of his state.

Cram, one of our new members, joins the faculty of Iowa State Teachers' College to work in the Extension Division in training the teachers in service, devoting most of his time to tests and measurements.

Staker at Aberdeen announces that he has been instrumental this past year in organizing the South Dakota Educational Research Association and is now its secretary.

Buckingham announces so many things we cannot print them all but one of special interest is the organization of a superintendents' conference to meet the first time this coming November.

We regret to announce the loss of another of our members. Mr. M. C. Lefler, formerly director of research, Lincoln, Nebraska, schools, has become sales manager of the University Publishing Co. of that city. We extend to him all good wishes in his new work.

It will be of interest to the members of our Association to learn, if they have not already learned, of the advancement of President Haggerty to the Deanship of the school of Education at the University of Minnesota. This deserved promotion came to Dr. Haggerty when Dr. Coffman advanced from the position of dean of education to the presidency of the same university. We may remark incidentally that we are likewise pleased that the presidency went to Dr. Coffman. What is more appropriate than that an educational man should be president of an educational institution?

We were about to record "with no small satisfaction" the appointment of Dean F. J. Kelly to the position of Associate Director of Educational and Mental Tests at the University of Michigan. We learned just in time, however, that Dean Kelly is to remain at Lawrence and assume an administrative position which had been newly created. He seems to be about to become adviser to the chancellor as well as to others in authority at the University of Kansas. Details from Dr. Kelly had not reached us at this writing. We know enough to be confident, however, that the position is an exceedingly important one.

Dr. Ballou, we are informed, has recently been elected city superintendent of schools at Washington, D. C. It seems to be impossible to keep good men at research work. Perhaps, however, as they rise to new heights they will be in a position to exalt research through the direct activities of others.

Now that you are back on the job again or soon will be and the summer play time is over, will you not drop the secretary a line telling us something you forgot before? And if you didn't do it before, remember every day may be the beginning of a new year. It is all up to you. The world is full of opportunities. What share of them are you going to seize?